GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 11, 2002, 13:46:28 ; Search time 71.25 Seconds
(without alignments)
886.220 Million cell updates/sec

Title: US-09-899-634A-4
Perfect score: 1898
Sequence: 1 MALLLCFVLLCGVADLTRSL.....SRMGAVPVMIPAQSKDGSIV 365

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1728.5	91.1	366	11	Q9DBJ8	Q9dbj8 mus musculu
2	1726	90.9	358	11	Q9R066	Q9r066 rattus norv
3	1657	87.3	319	6	Q9TU79	Q9tu79 sus scrofa
4	1633	86.0	344	11	Q9R067	Q9r067 rattus norv
5	1629	85.8	344	4	Q9UKV4	Q9ukv4 homo sapien
6	1613	85.0	352	11	Q91W66	Q91w66 mus musculu
7	1545	81.4	319	6	Q9TU80	Q9tu80 canis famil
8	1009	53.2	372	13	Q90Y50	Q90y50 brachydanio
9	641	33.8	164	11	Q99KG0	Q99kg0 mus musculu
10	471	24.8	373	4	Q9H6B4	Q9h6b4 homo sapien
11	440.5	23.2	373	11	Q920S5	Q920s5 mus musculu
12	360.5	19.0	394	11	Q925F2	Q925f2 mus musculu
13	336.5	17.7	319	11	Q922D5	Q922d5 mus musculu
14	335.5	17.7	390	6	Q95KI3	Q95ki3 macaca fasc
15	327.5	17.3	304	11	Q9CVA4	Q9cva4 mus musculu
16	326.5	17.2	319	11	Q9JKA5	Q9jka5 mus musculu
17	325.5	17.1	390	4	Q96AP7	Q96ap7 homo sapien
18	323.5	17.0	327	4	Q96IQ7	Q96iq7 homo sapien
19	323.5	17.0	390	4	Q96T50	Q96t50 homo sapien
20	322	17.0	335	13	Q9YGV5	Q9ygv5 gallus gall
21	319.5	16.8	325	4	O95791	O95791 homo sapien
22	317	16.7	318	13	Q91664	Q91664 xenopus lae
23	312	16.4	335	13	Q9PWR4	Q9pwr4 gallus gall
24	309.5	16.3	284	4	Q9NX42	Q9nx42 homo sapien
25	309	16.3	335	13	Q9YGH1	Q9ygh1 gallus gall
26	306	16.1	328	11	Q9Z109	Q9z109 mus musculu
27	290.5	15.3	407	11	Q9D2J4	Q9d2j4 mus musculu
28	284	15.0	248	11	Q9D0T4	Q9d0t4 mus musculu
29	258	13.6	259	4	O95532	O95532 homo sapien
30	224	11.8	300	11	Q9D9J0	Q9d9j0 mus musculu
31	223	11.7	300	11	Q9DA22	Q9da22 mus musculu
32	195.5	10.3	181	13	Q91665	Q91665 xenopus lae
33	174.5	9.2	300	11	Q9JHY1	Q9jhy1 rattus norv
34	172	9.1	280	13	O73716	O73716 grus americ
35	171.5	9.0	298	11	Q9JI59	Q9ji59 mus musculu
36	170	9.0	510	11	Q9JLB8	Q9jlb8 mus musculu
37	169.5	8.9	438	11	Q9JLB7	Q9jlb7 mus musculu
38	167	8.8	399	4	Q9Y279	Q9y279 homo sapien
39	166	8.7	259	4	Q9Y5B2	Q9y5b2 homo sapien
40	165	8.7	1094	4	Q9BYB8	Q9byb8 homo sapien
41	164.5	8.7	483	4	Q9UFI4	Q9ufi4 homo sapien
42	164.5	8.7	1093	4	Q96JA1	Q96ja1 homo sapien
43	162	8.5	740	11	Q9D332	Q9d332 mus musculu
44	161.5	8.5	306	11	Q9R1Z9	Q9r1z9 mus musculu
45	160.5	8.5	310	11	Q9EPK4	Q9epk4 mus musculu

ALIGNMENTS

```
RESULT 1
Q9DBJ8
ID   Q9DBJ8      PRELIMINARY;      PRT;   366 AA.
AC   Q9DBJ8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.
GN   CXADR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=LIVER;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK004908; BAB23660.1; -.
DR   MGD; MGI:1201679; Cxadr.
```

```
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain.
SQ   SEQUENCE   366 AA;  40078 MW;  D1C3CA7163F3B231 CRC64;


  Query Match             91.1%;  Score 1728.5;  DB 11;  Length 366;
  Best Local Similarity   90.2%;  Pred. No. 2.9e-145;
  Matches  330;  Conservative   17;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           || |||||||||:|| |  ||||||| |||||||||||||:||| ||||||||||||:|
Db       1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

        61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:||| |||||||||||||:|| ||||||||||||:|||||||||||||||||||||||
Db      61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| ||| ||||:|||||||:|||||||||||||||||||||| :||||||:
Db     121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           || :|| |||||||||||||||:||||||||:||||||||:||||||||||||||||||||
Db     181 SQTMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||| |||||||| |:|||:|||||||||||||||||||||||||||||:|||||
Db     241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPV-MIPAQS 359
           |||||||||||||||||||||||||||||||  |||||||||||||| ||||||
Db     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKVAAPNLSRMGAVPVMMIPAQS 360

Qy     360 KDGSIV 365
           ||||||
Db     361 KDGSIV 366


RESULT   2
Q9R066
ID   Q9R066      PRELIMINARY;      PRT;   358 AA.
AC   Q9R066;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIE-ADENOVIRUS-RECEPTOR HOMOLOG (FRAGMENT).
GN   CAR1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=LIVER;
RX   MEDLINE=99422053; PubMed=10490761;
RA   Fechner H., Haack A., Wang H., Wang X., Eizema K., Pauschinger M.,
RA   Schoemaker R.G., van Veghel R., Houtsmuller A.B., Schultheiss H.-P.,
RA   Lamers J.M.J., Poller W.;
RT   "Expression of Coxsackie-adenovirus-receptor and alpha v-integrin does
RT   not correlate with adenovector targeting in vivo indicating anatomical
RT   vector barriers.";
RL   Gene Ther. 6:1520-1535(1999).
DR   EMBL; AF109644; AAF01255.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Receptor.
FT   NON_TER     358    358
SQ   SEQUENCE   358 AA;  39261 MW;  731B7384A786BB04 CRC64;


  Query Match             90.9%;  Score 1726;  DB 11;  Length 358;
  Best Local Similarity   91.6%;  Pred. No. 4.6e-145;
  Matches  328;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||||| | ||||||||| |||||||||||||:||| ||||||||||||:|
Db       1 MALLLCFVLLCGVADFTSSLSITTPEQRIEKAKGETAYLPCKFTLEPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:|||||||||||||||||||:|| ||||||||||||:|||||||||||||||||||||||
Db      61 PSDNQKVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |:|  ||||:|||||||:|||| |||||||||||||||||| ||||||:
Db     121 KVKKAPGVANRKFLLTVLVKPSGTRCFVDGSGEIGNDFKLKCEPKEGSLPLQYEWQKLSD 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| |||||||||||||:||||||||:||||||||:||||||||||||||||||||
Db     181 SQKMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||| |||||||| |:|||||||||||||||||||||||||||||||:|||||
Db     241 GAVIGTLLALVLIGAILFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQ 358
           |||||||||||||||||||||||||||||||  |||||||||||||||||||
Db     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKVAAPNLSRMGAVPVMIPAQ 358


RESULT   3
Q9TU79
ID   Q9TU79      PRELIMINARY;      PRT;   319 AA.
AC   Q9TU79;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIE-ADENOVIRUS-RECEPTOR HOMOLOG (FRAGMENT).
GN   CAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=99422053; PubMed=10490761;
RA   Fechner H., Haack A., Wang H., Wang X., Eizema K., Pauschinger M.,
RA   Schoemaker R.G., van Veghel R., Houtsmuller A.B., Schultheiss H.-P.,
RA   Lamers J.M.J., Poller W.;
RT   "Expression of Coxsackie-adenovirus-receptor and alpha v-integrin does
RT   not correlate with adenovector targeting in vivo indicating anatomical
RT   vector barriers.";
RL   Gene Ther. 6:1520-1535(1999).
DR   EMBL; AF109646; AAF01257.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Receptor.
FT   NON_TER     319    319
SQ   SEQUENCE   319 AA;  35301 MW;  C485555A6C9F1B5B CRC64;
```

```
  Query Match           87.3%;  Score 1657;  DB 6;  Length 319;
  Best Local Similarity  99.7%;  Pred. No. 5.3e-139;
  Matches  318;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MALLLRFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYN 319
           |||||||||||||||||||
Db     301 SLGSMSPSNMEGYSKTQYN 319


RESULT   4
Q9R067
ID   Q9R067      PRELIMINARY;      PRT;   344 AA.
AC   Q9R067;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIE-ADENOVIRUS-RECEPTOR HOMOLOG (FRAGMENT).
GN   CAR2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=LIVER;
RX   MEDLINE=99422053; PubMed=10490761;
RA   Fechner H., Haack A., Wang H., Wang X., Eizema K., Pauschinger M.,
RA   Schoemaker R.G., van Veghel R., Houtsmuller A.B., Schultheiss H.-P.,
RA   Lamers J.M.J., Poller W.;
RT   "Expression of Coxsackie-adenovirus-receptor and alpha v-integrin does
RT   not correlate with adenovector targeting in vivo indicating anatomical
RT   vector barriers.";
RL   Gene Ther. 6:1520-1535(1999).
DR   EMBL; AF109643; AAF01254.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Receptor.
FT   NON_TER     344     344
SQ   SEQUENCE   344 AA;  38030 MW;  D436498820D8B9B7 CRC64;


  Query Match           86.0%;  Score 1633;  DB 11;  Length 344;
  Best Local Similarity  91.2%;  Pred. No. 8e-137;
  Matches  309;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           ||||||||||||||| | |||||||| ||||||||||||:||| ||||||||||||:|
Db       1 MALLLCFVLLCGVADFTSSLSITTPEQRIEKAKGETAYLPCKFTLEPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:|||||||||||||||||:|| |||||||||||:|||||||||||||||||||||||||
Db      61 PSDNQKVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |:|  ||||:|||||||:|||| |||||||||||||||| |||||||:
Db     121 KVKKAPGVANRKFLLTVLVKPSGTRCFVDGSGEIGNDFKLKCEPKEGSLPLQYEWQKLSD 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||||||||||:||||||||:|||||||:|||||||||||||||||||||
Db     181 SQKMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||| |||||||| |:|||||||||||||||||||||||||||||||||||:|||||
Db     241 GAVIGTLLALVLIGAILFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAK 339
           |||||||||||||||||||||||||||||||||||  ||
Db     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAK 339


RESULT   5
Q9UKV4
ID   Q9UKV4      PRELIMINARY;      PRT;   344 AA.
AC   Q9UKV4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIE AND ADENOVIRUS RECEPTOR PROTEIN (FRAGMENT).
GN   HCAR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=99422053; PubMed=10490761;
RA   Fechner H., Haack A., Wang H., Wang X., Eizema K., Pauschinger M.,
RA   Schoemaker R.G., van Veghel R., Houtsmuller A.B., Schultheiss H.-P.,
RA   Lamers J.M.J., Poller W.;
RT   "Expression of Coxsackie-adenovirus-receptor and alpha v-integrin does
RT   not correlate with adenovector targeting in vivo indicating anatomical
RT   vector barriers.";
RL   Gene Ther. 6:1520-1535(1999).
DR   EMBL; AF124598; AAD31772.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Receptor.
FT   NON_TER     344     344
SQ   SEQUENCE   344 AA;  38097 MW;  EEF3722E96C697AA CRC64;


  Query Match           85.8%;  Score 1629;  DB 4;  Length 344;
  Best Local Similarity  90.3%;  Pred. No. 1.8e-136;
  Matches  306;  Conservative   15;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db       1 MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| ||||||||||:|||:||||||||||| |||||||:
Db     121 KVKKAPGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSD 180
```

```
Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| ||||||| |||||||:|||||||||:||||||||||||:||||||:|| ||
Db    181 SQKMPTSWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||:||||||
Db    241 GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAK 339
          ||||||||||||||||||||||||||| ||||||| ||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAK 339


RESULT   6
Q91W66
ID   Q91W66      PRELIMINARY;      PRT;   352 AA.
AC   Q91W66;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC016457; AAH16457.1; -.
KW   Receptor.
SQ   SEQUENCE   352 AA;  38843 MW;  2BD8CBD25D8CE123 CRC64;


  Query Match             85.0%;  Score 1613;  DB 11;  Length 352;
  Best Local Similarity   89.7%;  Pred. No. 5e-135;
  Matches  304;  Conservative   17;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          || |||||||||:|| |  |||||||| ||||||||||||:||| ||||||||||||:|
Db      1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |:||| |||||||||||||:|| ||||||||||:|||||||||||||||||||||||||
Db     61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||  |||||:|||||||:|||||||||||||||||||| :|||||:
Db    121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          || :|| |||||||||||||||:||||||||:||||||||:||||||||||||||||||
Db    181 SQTMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||||| |||||||| |:||||:|||||||||||||||||||||||||||||:|||||
Db    241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAK 339
          |||||||||||||||||||||||||||||||||  ||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAK 339


RESULT   7
Q9TU80
ID   Q9TU80      PRELIMINARY;      PRT;   319 AA.
AC   Q9TU80;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIE-ADENOVIRUS-RECEPTOR HOMOLOG (FRAGMENT).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BREED GERMAN SHEPHERD; TISSUE=LIVER;
RX   MEDLINE=99422053; PubMed=10490761;
RA   Fechner H., Haack A., Wang H., Wang X., Eizema K., Pauschinger M.,
RA   Schoemaker R.G., van Veghel R., Houtsmuller A.B., Schultheiss H.-P.,
RA   Lamers J.M.J., Poller W.;
RT   "Expression of Coxsackie-adenovirus-receptor and alpha v-integrin does
RT   not correlate with adenovector targeting in vivo indicating anatomical
RT   vector barriers.";
RL   Gene Ther. 6:1520-1535(1999).
DR   EMBL; AF109645; AAF01256.1; -.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Receptor.
FT   NON_TER     319     319
SQ   SEQUENCE   319 AA;  35422 MW;  85C63A6EC7986965 CRC64;


  Query Match             81.4%;  Score 1545;  DB 6;  Length 319;
  Best Local Similarity   92.2%;  Pred. No. 4.8e-129;
  Matches  294;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          ||||| |||||||||| ||||||||||||||||||||||||:||| ||||||||||||||
Db      1 MALLLRFVLLCGVADFTRSLSITTPEQMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLLS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLRLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          ||||||||||||||||||:|||| |||||||||||||||||||||||| ||||||||
Db    121 KVKKAPGVGNKKIQLTVLVKPSGIRCYVDGSEEIGNDFKLKCEPKEGSLPLQYEWQKLSN 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          ||| |  |  :|||||||:|||||||||| ||| |||||||| |||:||||||||||||
Db    181 SQKTPPSWSTDMTSPVISIKNASTEYSGTYICTVTNRVGSDQCQLRLNVVPPSNRAGTIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          |||:|:|||||||| ||||| ||||||||||||||||||| ||||||||||:|||||
Db    241 GAVVGILLALVLIGFIVFCCRKKRREEKYEKEVHHDIREDVPPLKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYN 319
          |||||||||||||||||||
Db    301 SLGSMSPSNMEGYSKTQYN 319


RESULT   8
Q90Y50
ID   Q90Y50      PRELIMINARY;      PRT;   372 AA.
AC   Q90Y50;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR-LIKE PROTEIN.
GN   CAR.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20535037; PubMed=11080637;
RA   van Raaij M.J., Chouin E., van der Zandt H., Bergelson J.M.,
```

```
RA   Cusack S.;
RT   "Dimeric structure of the coxsackievirus and adenovirus receptor D1
RT   domain at 1.7 A resolution.";
RL   Structure 8:1147-1155(2000).
DR   EMBL; AF268197; AAK58592.1; -.
KW   Receptor.
SQ   SEQUENCE   372 AA;  40664 MW;  C363B71E7601C73A CRC64;


  Query Match           53.2%;  Score 1009;  DB 13;  Length 372;
  Best Local Similarity   50.8%;  Pred. No. 2.6e-81;
  Matches  189;  Conservative   78;  Mismatches   89;  Indels   16;  Gaps    5;

Qy       3 LLLCFVLL-----CGVADLTRSLSIT-TPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIE 56
           | : :|:|     ||       | || | :  |||| ||:  | |:|||  :| ||||||
Db       8 LCVTYVILLTGSACG-------LQITSTGQTSIEKASGESVKLDCQFTLASDDSGPLDIE 60

Qy      57 WLLSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIG 116
           | | |:|||| ::|:|:|||||: :: ||  ||||||| | | |:||||:|:  |: :| |
Db      61 WSLQPSDNQKEEKVVIVYSGDRAFEHYYDPLKGRVHFNSPDPKNGDASMNIMGLKATDTG 120

Qy     117 TYQCKVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQ 176
           |||||:|| ||: ::|  |||:::||  :|  :|   :| :  |||    ||: |: | |:
Db     121 TYQCKIKKVPGIASRKYLLTVMVRPSKPKCSAEGQTYVGKNMVLKCSSVEGTQPMEYIWE 180

Qy     177 KLSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRA 236
           : | :: || | : :  :  :::|||: : |||| |  |||||:::|:: : :  | | |
Db     181 RTSGNKLLPPLAILDKVTGTMTLKNATGDASGTYRCQAKNRVGTEECVVEVTITQPPNTA 240

Qy     237 GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSY-- 294
           | ||| :| :|| |:|: ||:||| : | ::|||||: ::|||||||||| |||||:
Db     241 GIIAGVIICILLLLILLALILFCCCRARHKKKYEKEIAYEIREDVPPPKSRVSTARSFTS 300

Qy     295 LGSNHSSLGSMSPSNMEGYSKTQYNQVPSEDFERAP-QSPTLPLAKVAAPNLSRMGAVPV 353
           :||  |||||||||||:  ||| ||:::|||:::| |  :|  | :::| ||||||||:||
Db     301 VGSQRSSLGSMSPSNLHEYSKPQYDKIPSEEYDRPPSHAPIPPPSRMAGPNLSRMGAIPV 360

Qy     354 MIPAQSKDGSIV 365
           |||||:||||||
Db     361 MIPAQNKDGSIV 372


RESULT   9
Q99KG0
ID   Q99KG0      PRELIMINARY;      PRT;   164 AA.
AC   Q99KG0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   SIMILAR TO COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC004680; AAH04680.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00409; IG; 1.
KW   Receptor.
SQ   SEQUENCE   164 AA;  17781 MW;  9DEFCBD456240C73 CRC64;


  Query Match           33.8%;  Score 641;  DB 11;  Length 164;
  Best Local Similarity   86.6%;  Pred. No. 4e-49;
  Matches  123;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           || |||||||||:|| |  |||||||| |||||||||||||:||| ||||||||||||:|

Db       1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:||| |||||||||||||:|| ||||||||||||:|||||||||||||||||||||||
Db      61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPS 142
           |||||||| |||  |||| | |
Db     121 KVKKAPGVANKKFLLTVLGKSS 142


RESULT  10
Q9H6B4
ID   Q9H6B4      PRELIMINARY;      PRT;   373 AA.
AC   Q9H6B4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   CDNA: FLJ22415 FIS, CLONE HRC08561 (HYPOTHETICAL 41.3 KDA
DE   PROTEIN).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA   Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MUSCLE, RHABDOMYOSARCOMA;.
RA   Strausberg R.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK026068; BAB15347.1; -.
DR   EMBL; BC009371; AAH09371.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Hypothetical protein.
SQ   SEQUENCE   373 AA;  41280 MW;  FDA215EB3B3C4335 CRC64;


  Query Match           24.8%;  Score 471;  DB 4;  Length 373;
  Best Local Similarity   32.6%;  Pred. No. 1.7e-33;
  Matches  124;  Conservative   70;  Mismatches  154;  Indels   32;  Gaps   10;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |:||| :|   |:    :|   |    |::   |   |||   ||  ::  |||||||:
Db       1 MSLLLLLLL---VSYYVGTLGTHTE---IKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLT 54

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
             ||:   :|:| ||   :|::  :: |||| | || | :||||: :  |: || | | |
Db      55 --DNEGNQKVVITYSSRHVYNNLTEEQKGRVAFASNFL-AGDASLQIEPLKPSDEGRYTC 111

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKL-- 178
           |||  :       : | ||::||  :| ::|    |:|  |:||   |: |::| ||::
Db     112 KVKNSGRYVWSHVILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIRE 171

Qy     179 --SNSQKLPTLWLAEMTSP-VISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNR 235
               ::||     :   |  : ::| :  ||| | ||  |  | : |::|:  |
Db     172 KEGEDERLPPKSRIDYNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRV-TVQYVQS 230

Qy     236 AGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVH-HDIREDVPPPKSRTSTARSY 294
           | :|||| |::   :|| |:|:   ::: :|:||:|   ::|||   ||:|.    |
Db     231 IGMVAGAVTGIVAGALLIFLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSS 290
```

```
Qy    295 LGSNHSSLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLA----------KV 340
Db    291 SSGSRSSRSGSSSTRSTANSASRSQRTLSTD--AAPQPGLATQAYSLVGPEVRGSEPKKV 348

Qy    341 AAPNLSRMGAVPVMIPAQSK 360
Db    349 HHANLTKAETTPSMIPSQSR 368
```

RESULT 11
Q920S5
ID Q920S5 PRELIMINARY; PRT; 373 AA.
AC Q920S5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ADIPOCYTE-SPECIFIC PROTEIN 5.
GN ASP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tsuruga H.;
RT "Adipocyte-specific protein 5, a novel protein upregulated during
RT adipocyte differentiation.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB040490; BAB68503.1; -.
SQ SEQUENCE 373 AA; 41186 MW; 5C6280584AF95326 CRC64;

Query Match 23.2%; Score 440.5; DB 11; Length 373;
Best Local Similarity 31.7%; Pred. No. 8.5e-31;
Matches 112; Conservative 68; Mismatches 146; Indels 27; Gaps 9;

```
Qy     29 IEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQVIILYSGDKIYDDYYQDLK 88
Db     22 IKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLT--DNEGNQKVVITYSSRHVYNNLTEEQK 79

Qy     89 GRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLLKPSGTRCYV 148
Db     80 GRVAFASNFL-AGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVILKALVRPSKPKCEL 138

Qy    149 DGSEEIGNDFKLKCEPKEGSLPLLYEWQKL----SNSQKLPTLWLAEMTSP-VISVKNAS 203
Db    139 EGEPTEGSDLTLQCESASGTKPIVYYWQRIREKEGEDEHLPPKSRIDYNNPGRVLLQNLT 198

Qy    204 TEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVFCCHKK 263
Db    199 MASSGLYQCTAGNEAGKESCVVRV-TVQYVQSIGMVAGAVTGIVAGALLIFLLIWLLIRR 257

Qy    264 RREEKYEKEVH-HDIREDVPPPKSRTSTARSYLGSNHSSLGSMSPSNMEGYSKTQYNQVP 322
Db    258 KSKDRYEEEDRPNEIREDAEAPRARLVKPSSSSSGSRSSRSGSSSTRSTGNSASRSQRTL 317

Qy    323 SEDFERAPQSPTL-PLA----------KVAAPNLSRMGAVPVMIPAQSK 360
Db    318 SS--EAAPQQPGLAPQAYSLIGPEVRGSEPKKVHHTTLTKAETTLSTTPSQSK 368
```

RESULT 12
Q925F2
ID Q925F2 PRELIMINARY; PRT; 394 AA.
AC Q925F2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIAL CELL-SELECTIVE ADHESION MOLECULE.
GN ESAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER/NIH;
RX MEDLINE=21238298; PubMed=11279107;
RA Hirata K.-I., Ishida T., Penta K., Rezaee M., Yang E., Wohlgemuth J.,
RA Quertermous T.;
RT "Cloning of an immunoglobulin family adhesion molecule selectively
RT expressed by endothelial cells.";
RL J. Biol. Chem. 276:16223-16231(2001).
DR EMBL; AF361882; AAK51504.1; -.
SQ SEQUENCE 394 AA; 41810 MW; 3D2B354943A2227D CRC64;

Query Match 19.0%; Score 360.5; DB 11; Length 394;
Best Local Similarity 28.9%; Pred. No. 1.2e-23;
Matches 116; Conservative 73; Mismatches 154; Indels 59; Gaps 15;

```
Qy      3 LLLCFVLLCGVADLTRS---LSITTPEQMIEKAKGETAYLPCRFTLGPED--QGPLDIE- 56
Db     13 LRVLFLGLSTLAAFSRAQMELHVPPGLNKLEAVEGEEVVLPAWYTMAREESWSHPREVPI 72

Qy     57 --WLLSPADNQKVDQVIILYSGDKIYDDYYQDLKGR--VHFTSNDLKSGDASINVTNLQL 112
Db     73 LIWFLE-QEGKEPNQVLSYING------VMTNKPGTALVH----SISSRNVSLRLGALQE 121

Qy    113 SDIGTYQCKVK------KAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKE 166
Db    122 GDSGTYRCSVNVQNDEGKSIGHSIKSIELKVLVPPAPPSCSLQGVPYVGTNVTLNCKSPR 181

Qy    167 GSLPLLYEWQKLSNSQKL---PTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQC 223
Db    182 SKPTAQYQWERLAPSSQVFFGPAL---DAVRGSLKLTNLSIAMSGVYVCKAQNRVGFAKC 238

Qy    224 LLRLDVVPPSNRAGTIAGAVIGVLLALVLI-GLIVFCCHKKRREEKYEKEVHHDIREDVP 282
Db    239 NVTLDVMTGS-KAAVVAGAVVGTFVGLVLIAGLVLL----YQRRSKTLEELANDIKEDAI 293

Qy    283 PPKSRTSTARSYLGSNHSSL------------------GSMSPSNMEGYSKTQYNQVPS 323
Db    294 APRTLPWTKGSDTISKNGTLSSVTSARALRPPKAAPPRPGTFTPTPSVSSQALSSPRLPR 353

Qy    324 EDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSKDGSIV 365
Db    354 VD-EPPPQAVSLTPGGVSSSALSRMGAVPVMVPAQSQAGSLV 394
```

RESULT 13
Q922D5
ID Q922D5 PRELIMINARY; PRT; 319 AA.
AC Q922D5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOPROTEIN A33 (TRANSMEMBRANE).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC008528; AAH08528.1; -.
KW Transmembrane.
SQ SEQUENCE 319 AA; 35692 MW; A37C8DB4BF685C3A CRC64;

Query Match 17.7%; Score 336.5; DB 11; Length 319;
Best Local Similarity 30.3%; Pred. No. 1.2e-21;
Matches 88; Conservative 56; Mismatches 113; Indels 33; Gaps 11;

```
Qy       9 LLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW--LLSPADNQK 66
           :|| :     :|:: | : ::  |:| :  ||| :     |:    |:|  ||      :
Db      11 MLCAIWVAADALTVETTQDILRAARGRSVTLPCTYNTYVSDREGF-IQWDKLL----RSQ 65

Qy      67 VDQVII--LYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK 124
            ::|:     :   || : |::   ||   ||| :  :||| :  | : | |||:| |
Db      66 TERVVTWNFVTKKYIYGNRYEN---RVR-VSNDAELSNASITIDQLTMDDNGTYECSVSL 121

Qy     125 APGV---GNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKL-SN 180
               :::| ||: ||   | : |   |||: :| |   |||    | |:   :
Db     122 MSDQDVNAKSRVRLLVLVPPSKPDCSIQGEMVIGNNIQLTCHSAEGSPSPQYSWKSYNAQ 181

Qy     181 SQKLPTLWLAEMTSPV----ISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRA 236
           :|: |      :| ||    : :|| ||| :| | ||  | || : | : :   |||
Db     182 NQQRP------LTQPVSGEPLLLKNISTETAGYYICTSSNDVGIESCNITVAPRPPSMNI 235

Qy     237 GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKS 286
             ||   || :||::||:||:||  : :::| :       |||  | ::
Db     236 ALYAGIAGGVFVALIIIGVIVYCCCCREKDDKDQD------REDARPNRA 279


RESULT  14
Q95KI3
ID   Q95KI3      PRELIMINARY;      PRT;   390 AA.
AC   Q95KI3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 40.9 KDA PROTEIN.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TEMPORAL LOBE RIGHT;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB060855; BAB46874.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   390 AA;  40946 MW;  CDBF63F2BD464EF5 CRC64;


  Query Match             17.7%;  Score 335.5;  DB 6;  Length 390;
  Best Local Similarity   27.2%;  Pred. No. 1.9e-21;
  Matches  111;  Conservative   69;  Mismatches  153;  Indels   75;  Gaps   15;

Qy       3 LLLCFVLLCGVADLTRS-LSITTPEQMIEKAKGETAYLPCRFTLGPE----DQGPLD-IE 56
           |   |: |  :|  :|: | :  |   ::  :|    ||  :||  |      | :  :
Db      13 LRFLFLGLSALAPPSRAELQLHLPANQLQAVEGGEVVLPAWYTLHAEVSSAQPGEVPFVM 72

Qy      57 WLLSPADNQKVDQVIILYSG--------DKIYDDYYQDLKGRVHFTSNDLKSGDASINVT 108
           |      | :| |||:   :|          :|                 : | : |: :
Db      73 WFFK--DKEKEDQVLSYINGVTTSKPGVSLVY----------------SMPSRNLSLRLE 114

Qy     109 NLQLSDIGTYQCKVK------KAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKC 162
            ||  | | | | |       :| |   | ::| ||: |:   | : |   :| :  | |
Db     115 GLQEKDSGPYSCSVNVQDKNGQASGHSIKTLELNVLVPPAPPSCRLQGVPRVGANVTLSC 174

Qy     163 EPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPV-------ISVKNASTEYSGTYSCTVK 215
           :       : |:| :     :||:       :||       :|: | |:  :| | |
Db     175 QSPRSKPAVQYQWDR-----QLPS--FQTFFAPVLDVIRGSLSLTNLSSSMAGVYVCKAH 227

Qy     216 NRVGSDQCLLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHH 275
           | ||: || : |: |     |  :||||:| |: | |:  :|   |   |  |  :|  :
Db     228 NEVGTAQCNVTLE-VSTGPGAAVVAGAVVGTLVGLGLLAGLVLLYH---RRGKALEEPAN 283

Qy     276 DIRED------VPPPKSRTSTARSYLGSNHSSL-----------GSMSPSNMEGYSKTQ 317
           ||:||      :| |||  : :::   |: :|             |:::|:
```

```
Db     284 DIKEDAIAPRTLPWPKSSDTISKNGTLSSVTSARALRPPHGPPRPGALTPTPSLSSQALP 343

Qy     318 YNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSKDGSIV 365
             ::|: |    ||  :|    |::  ||||||||||:||||: ||:|
Db     344 SPRLPTTDGAN-PQPISLIPGGVSSSGLSRMGAVPVMVPAQSQAGSLV 390


RESULT  15
Q9CVA4
ID   Q9CVA4      PRELIMINARY;      PRT;   304 AA.
AC   Q9CVA4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   2210413P10RIK PROTEIN (FRAGMENT).
GN   2210413P10RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=STOMACH;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK008920; BAB25968.1; -.
DR   MGD; MGI:1919617; 2210413P10Rik.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00408; IGc2; 2.
DR   SMART; SM00406; IGv; 1.
KW   Immunoglobulin domain.
FT   NON_TER     304     304
SQ   SEQUENCE   304 AA;  31919 MW;  522BA38898AD7A9F CRC64;


  Query Match             17.3%;  Score 327.5;  DB 11;  Length 304;
  Best Local Similarity   31.3%;  Pred. No. 7e-21;
  Matches   91;  Conservative   45;  Mismatches  132;  Indels   23;  Gaps   10;

Qy       4 LLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW-LLSPA 62
           || || | |:|     : :| | : :   ||:|| | | :     |     :||  : |
Db      14 LLGFVCLSGLA-----VEVTVPTEPLSVPKGKTAELSCSYKTSVGDN--FALEWSFVQPG 66

Qy      63 DNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKV 122
                 :: ::   :|         |     :    | |:: :|:|: || ||| | |
Db      67 KPISASVPVLYFTNGHLYPT--GSKADRAILLHDPPTGGLATLKLTDLRPSDTGTYLCNV 124

Qy     123 KKAP-----GVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQK 177
               |     |:|   | ||||: ||   |   |   :|    | |   ||:   :| |::
```

```
Db    125 NNPPDFYTNGLG--LINLTVLVPPSHPLCSQSGQTSVGGSAALGCRSSEGAPKPVYNWER 182

Qy    178 LSNSQKLPT-LWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRA 236
          | :|   |    : :  |  : : | |   ||| |   |::||  | | | |   |
Db    183 LGSSPTPPPGSMVQDEVSGQLILTNLSLTSSGTYRCVASNQMGSASCELNLSVTDSSE-- 240

Qy    237 GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYE-KEVH--HDIREDVPPP 284
          | :|| :||||| ::|: :  ||  : ::| | | || :   |:|||   |
Db    241 GRVAGTLIGVLLGVLLLSVAAFCLIRFQKERKKEPKETYGGSDLREDATAP 291


Search completed: May 11, 2002, 13:50:13
Job time: 225 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 11, 2002, 13:46:33 ; Search time 28.31 Seconds
(without alignments)
499.210 Million cell updates/sec

Title: US-09-899-634A-4
Perfect score: 1898
Sequence: 1 MALLLCFVLLCGVADLTRSL.....SRMGAVPVMIPAQSKDGSIV 365

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1754	92.4	365	1	CXAR_HUMAN	P78310 homo sapien
2	1739	91.6	365	1	CXAR_MOUSE	P97792 mus musculu
3	368.5	19.4	319	1	A33_HUMAN	Q99795 homo sapien
4	197	10.4	299	1	JAM1_HUMAN	Q9y624 homo sapien
5	186	9.8	298	1	JAM2_HUMAN	P57087 homo sapien
6	170	9.0	298	1	JAM1_BOVIN	Q9xt56 bos taurus
7	161.5	8.5	306	1	CD80_MOUSE	Q00609 mus musculu
8	160.5	8.5	300	1	JAM1_MOUSE	O88792 mus musculu
9	152.5	8.0	538	1	PVR2_HUMAN	Q92692 homo sapien
10	151	8.0	526	1	CEA1_HUMAN	P13688 homo sapien
11	149	7.9	740	1	PEC1_PIG	Q95242 sus scrofa
12	148.5	7.8	727	1	PEC1_MOUSE	Q08481 mus musculu
13	148.5	7.8	764	1	ICCR_DROME	Q08180 drosophila
14	148	7.8	526	1	BUTY_BOVIN	P18892 bos taurus
15	146.5	7.7	519	1	ECTO_RAT	P16573 rattus norv
16	145.5	7.7	517	1	PVR1_HUMAN	Q15223 homo sapien
17	139.5	7.3	738	1	PEC1_HUMAN	P16284 homo sapien
18	137.5	7.2	515	1	PVR1_PIG	Q9gl76 sus scrofa
19	136.5	7.2	3707	1	PGBM_MOUSE	Q05793 mus musculu
20	136	7.2	309	1	CD86_MOUSE	P42082 mus musculu
21	134	7.1	555	1	C166_CARAU	Q90304 carassius a
22	133	7.0	329	1	CD86_HUMAN	P42081 homo sapien
23	133	7.0	526	1	BUTY_HUMAN	Q13410 homo sapien
24	133	7.0	530	1	PVR2_MOUSE	P32507 mus musculu
25	132	7.0	515	1	PVR1_MOUSE	Q9jkf6 mus musculu
26	131.5	6.9	348	1	KILO_RAT	Q9z0j8 rattus norv
27	130.5	6.9	1070	1	PTK7_HUMAN	Q13308 homo sapien
28	130	6.8	837	1	NCM2_MOUSE	O35136 mus musculu
29	128.5	6.8	4393	1	PGBM_HUMAN	P98160 homo sapien
30	127.5	6.7	344	1	NTRI_RAT	Q62718 rattus norv
31	127	6.7	246	1	MYP0_HETFR	P20938 heterodontu
32	126.5	6.7	564	1	C166_BRARE	Q90460 brachydanio
33	125.5	6.6	739	1	VCA1_HUMAN	P19320 homo sapien
34	124.5	6.6	847	1	CD22_HUMAN	P20273 homo sapien
35	124.5	6.6	858	1	NCA1_RAT	P13596 rattus norv
36	124	6.5	521	1	CEA1_MOUSE	P31809 mus musculu
37	124	6.5	1336	1	VGR1_RAT	P53767 rattus norv
38	123.5	6.5	417	1	PVR_CERAE	P32506 cercopithec
39	123	6.5	1906	1	KMLS_CHICK	P11799 gallus gall
40	122.5	6.5	811	1	FS22_DROME	P34083 drosophila
41	122.5	6.5	873	1	FS21_DROME	P34082 drosophila
42	122	6.4	215	1	CIB2_RAT	P54900 rattus norv
43	122	6.4	837	1	NCM2_HUMAN	O15394 homo sapien
44	122	6.4	1036	1	AXO1_CHICK	P28685 gallus gall
45	121	6.4	219	1	MYP0_BOVIN	P10522 bos taurus

ALIGNMENTS

RESULT 1
CXAR_HUMAN
ID CXAR_HUMAN STANDARD; PRT; 365 AA.
AC P78310; O00694;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Coxsackievirus and adenovirus receptor precursor (Coxsackievirus B-
DE adenovirus receptor) (hCAR) (CVB3 binding protein).
GN CXADR OR CAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97190109; PubMed=9036860;
RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,
RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;
RT "Isolation of a common receptor for Coxsackie B viruses and
RT adenoviruses 2 and 5.";
RL Science 275:1320-1323(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97250541; PubMed=9096397;
RA Tomko R.P., Xu R., Philipson L.;
RT "HCAR and MCAR: the human and mouse cellular receptors for subgroup C
RT adenoviruses and group B coxsackieviruses.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20008750; PubMed=10543405;
RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,
RA Bowles N.E.;
RT "Genomic organization and chromosomal localization of the human
RT Coxsackievirus B-adenovirus receptor gene.";
RL Hum. Genet. 105:354-359(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA Anderson C.W., Kieleczawa J., Dunn J.J., Freimuth P.;
RT "Sequence and expression of CXADR, the human gene for the
RT coxsackievirus and adenovirus receptor.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND
CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC ---

```
CC   ---------------------------------------------------------------------
DR   EMBL; Y07593; CAA68868.1; -.
DR   EMBL; U90716; AAC51234.1; -.
DR   EMBL; AF169366; AAF05908.1; -.
DR   EMBL; AF169360; AAF05908.1; JOINED.
DR   EMBL; AF169361; AAF05908.1; JOINED.
DR   EMBL; AF169362; AAF05908.1; JOINED.
DR   EMBL; AF169363; AAF05908.1; JOINED.
DR   EMBL; AF169364; AAF05908.1; JOINED.
DR   EMBL; AF169365; AAF05908.1; JOINED.
DR   EMBL; AF200465; AAF24344.1; -.
DR   MIM; 602621; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
KW   Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;
     Repeat.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    365       COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.
FT   DOMAIN       20    237       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    238    258       POTENTIAL.
FT   DOMAIN      259    365       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       34    127       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      155    219       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID     41    120       BY SIMILARITY.
FT   DISULFID    162    212       BY SIMILARITY.
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    201    201       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   365 AA;  40029 MW;  AB01C6346CB7FE64 CRC64;


  Query Match             92.4%;  Score 1754;  DB 1;  Length 365;
  Best Local Similarity   90.7%;  Pred. No. 3.1e-126;
  Matches  331;  Conservative   16;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           ||||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db       1 MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      61 PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db     121 KVKKAPGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSD 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||| |||||||:||||||||||:||||||||||||:|||||:|| ||
Db     181 SQKMPTSWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||:|| |||| |||||:||| ||||||||||||||||||||||||||||:|||||
Db     241 GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
           |||||||||||||||||||||||||| ||||||| |||||||||||||:|||||||||
Db     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAKVAAPNLSRMGAIPVMIPAQSK 360

Qy     361 DGSIV 365
           |||||
Db     361 DGSIV 365


RESULT   2
CXAR_MOUSE
ID   CXAR_MOUSE     STANDARD;      PRT;   365 AA.
AC   P97792; O09052;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Coxsackievirus and adenovirus receptor homolog precursor (mCAR).
GN   CXADR OR CAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=97190109; PubMed=9036860;
RA   Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,
RA   Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;
RT   "Isolation of a common receptor for Coxsackie B viruses and
RT   adenoviruses 2 and 5.";
RL   Science 275:1320-1323(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/MAI;
RX   MEDLINE=97250541; PubMed=9096397;
RA   Tomko R.P., Xu R., Philipson L.;
RT   "HCAR and MCAR: the human and mouse cellular receptors for subgroup C
RT   adenoviruses and group B coxsackieviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RA   Bergelson J.M., Krithivas A., Crowell T.L., Finberg R.W.;
RT   "The murine CAR homologue (mCAR) is a receptor for coxsackie B
RT   viruses and adenoviruses.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y10320; CAA71368.1; -.
DR   EMBL; U90715; AAC53148.1; -.
DR   EMBL; Y11929; CAA72679.1; -.
DR   MGD; MGI:1201679; Cxadr.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
KW   Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;
KW   Repeat.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    365       COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR
FT                                HOMOLOG.
FT   DOMAIN       20    237       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    238    258       POTENTIAL.
FT   DOMAIN      259    365       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       34    127       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      155    219       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID     41    120       BY SIMILARITY.
FT   DISULFID    162    212       BY SIMILARITY.
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    201    201       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    340    365       VAAPNLSRMGAVPVMIPAQSKDGSIV -> FKYAYKTDGIT
FT                                VV (IN REF. 2 AND 3).
SQ   SEQUENCE   365 AA;  39947 MW;  5445B4B52A34B2A2 CRC64;


  Query Match             91.6%;  Score 1739;  DB 1;  Length 365;
```

```
  Best Local Similarity   90.4%;  Pred. No. 4.2e-125;
  Matches  330;  Conservative   17;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           || |||||||||:|| |  |||||||| |||||||||||||:||| ||||||||||||:|
Db       1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:||| |||||||||||||||:|| ||||||||||||:|||||||||||||||||||||||
Db      61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||  |||||:|||||||:|||||||||||||||||||||| :||||||:
Db     121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           || :|| |||||||||||||||:|||||||||:||||||||:|||||||||||||||||||
Db     181 SQTMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||| |||||||| |:||||:|||||||||||||||||||||||||||||:|||||
Db     241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
           |||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKVAAPNLSRMGAVPVMIPAQSK 360

Qy     361 DGSIV 365
           |||||
Db     361 DGSIV 365
```

```
RESULT   3
A33_HUMAN
ID   A33_HUMAN      STANDARD;      PRT;   319 AA.
AC   Q99795;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Cell surface A33 antigen precursor (Glycoprotein A33).
GN   GPA33.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Colon carcinoma;
RX   MEDLINE=97165045; PubMed=9012807;
RA   Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,
RA   Moritz R.L., Tu G.-F., Ji H., Whitehead R.H., Groenen L.C.,
RA   Scott A.M., Ritter G., Cohen L., Welt S., Old L.J., Nice E.C.,
RA   Burgess A.W.;
RT   "The human A33 antigen is a transmembrane glycoprotein and a novel
RT   member of the immunoglobulin superfamily.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).
RN   [2]
RP   POST-TRANSLATIONAL MODIFICATIONS.
RX   MEDLINE=97396159; PubMed=9245713;
RA   Ritter G., Cohen L.S., Nice E.C., Catimel B., Burgess A.W.,
RA   Moritz R.L., Ji H., Heath J.K., White S.J., Welt S., Old L.J.,
RA   Simpson R.J.;
RT   "Characterization of posttranslational modifications of human A33
RT   antigen, a novel palmitoylated surface glycoprotein of human
RT   gastrointestinal epithelium.";
RL   Biochem. Biophys. Res. Commun. 236:682-686(1997).
CC   -!- FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL
CC       EPITHELIUM AND IN 95% OF COLON CANCERS.
CC   -!- PTM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 K OF N-LINKED
CC       CARBOHYDRATE.
CC   -!- PTM: PALMITOYLATED.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U79725; AAC50957.1; -.
DR   HSSP; P06907; 1NEU.
DR   MIM; 602171; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00406; IGv; 1.
KW   Immunoglobulin domain; Lipoprotein; Palmitate; Glycoprotein;
KW   Transmembrane; Signal; Antigen.
FT   SIGNAL        1     21
FT   CHAIN        22    319       CELL SURFACE A33 ANTIGEN.
FT   DOMAIN       22    235       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    236    256       POTENTIAL.
FT   DOMAIN      257    319       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       36    124       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      139    229       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      258    261       POLY-CYS.
FT   DISULFID     43    117       POTENTIAL.
FT   DISULFID    162    211       POTENTIAL.
FT   DISULFID    146    222       POTENTIAL.
FT   CARBOHYD    112    112       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    200    200       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   319 AA;  35632 MW;  9BFC7AAF45C2408E CRC64;
```

```
  Query Match           19.4%;  Score 368.5;  DB 1;  Length 319;
  Best Local Similarity   30.9%;  Pred. No. 5.9e-21;
  Matches   99;  Conservative   62;  Mismatches  128;  Indels   31;  Gaps   13;

Qy      10 LCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW---LLSPADNQK 66
           || |      ::|: ||: ::  ::|::  ||| :       :  | |:|   ||:
Db      12 LCAVRVTVDAISVETPQDVLRASQGKSVTLPCTYHTSTSSREGL-IQWDKLLLT-----H 65

Qy      67 VDQVIILYSGDK--IYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK 124
           ::|:|    :|  |: : |   | ||   ||: :  |||| :  | ::| |||:| |
Db      66 TERVVIWPFSNKNYIHGELY---KNRVSI-SNNAEQSDASITIDQLTMADNGTYECSVSL 121

Qy     125 APGV-GNKK--IQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLS-N 180
           : || |  ::| ||: ||   | ::|   |||: :| |: ||||    | |:: :
Db     122 MSDLEGNTKSRVRLLVLVPPSKPECGIEGETIIGNNIQLTCQSKEGSPTPQYSWKRYNIL 181

Qy     181 SQKLPTLWLAEMTS-PVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTI 239
           :|: |   ||:  |   :|:|| ||: || | ||  |  |:  | : : |  ||
Db     182 NQEQP---LAQPASGQPVSLKNISTDTSGYYICTSSNEEGTQFCNITVAVRSPSMNVALY 238

Qy     240 AGAVIGVLLALVLIGLIVFCC----HKKRREEKYEKEVHHDIREDVPPPKSRTSTAR--- 292
           |  :||: ||::||:|::||         |:| :   : :  |: |      |  |
Db     239 VGIAVGVVAALIIIGIIIYCCCCRGKDDNTEDKEDARPNREAYEEPPEQLRELSREREEE 298

Qy     293 -SYLGSNHSSLGSMSPSNME 311
           |      | |  || :::
Db     299 DDYRQEEQRSTGRESPDHLD 318
```

```
RESULT   4
JAM1_HUMAN
ID   JAM1_HUMAN      STANDARD;      PRT;   299 AA.
AC   Q9Y624;
```

```
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Junctional adhesion molecule 1 precursor (JAM) (Platelet adhesion
DE   molecule 1) (PAM-1) (Platelet F11 receptor).
GN   JAM1 OR JCAM.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99323940; PubMed=10395639;
RA   Ozaki H., Ishii K., Horiuchi H., Arai H., Kawamoto T., Okawa K.,
RA   Iwamatsu A., Kita T.;
RT   "Cutting edge: combined treatment of TNF-alpha and IFN-gamma causes
RT   redistribution of junctional adhesion molecule in human endothelial
RT   cells.";
RL   J. Immunol. 163:553-557(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sobocka M.B., Sobocki T., Rushbrook J.I., Banerjee P., Weiss C.,
RA   Kornecki E.;
RT   "Molecular cloning and sequencing of the cDNA of F11 receptor, a
RT   novel Ig superfamily member from human platelets.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Naik U.P., Naik M.U., DeLeon P., Spychala J.;
RT   "Cloning and characterization of PAM-1, a novel platelet adhesion
RT   molecule involved in platelet activation.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a catalog of human genes and proteins: sequencing and
RT   analysis of 500 novel complete protein coding human cDNAs.";
RL   Genome Res. 11:422-435(2001).
CC   -!- FUNCTION: PLAYS A ROLE IN REGULATING MONOCYTE TRANSMIGRATION
CC       INVOLVED IN INTEGRITY OF EPITHELIAL BARRIER. INVOLVED IN PLATELET
CC       ACTIVATION.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: LOCALIZED AT TIGHT JUNCTIONS OF BOTH
CC       EPITHELIAL AND ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF111713; AAD42050.1; -.
DR   EMBL; AF207907; AAF22829.1; -.
DR   EMBL; AF172398; AAD48877.1; -.
DR   EMBL; AL136649; CAB66584.1; -.
DR   MIM; 605721; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00406; IGv; 1.
KW   Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW   Repeat; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    299       JUNCTIONAL ADHESION MOLECULE 1.
FT   DOMAIN       26    238       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    239    259       POTENTIAL.
FT   DOMAIN      260    299       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       43    116       IG-LIKE V-TYPE DOMAIN 1.
FT   DOMAIN      146    219       IG-LIKE V-TYPE DOMAIN 2.
FT   DISULFID     50    109       POTENTIAL.
FT   DISULFID    153    212       POTENTIAL.
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   299 AA;  32583 MW;  D95DE2FEA23D2851 CRC64;


  Query Match             10.4%;  Score 197;  DB 1;  Length 299;
  Best Local Similarity   26.3%;  Pred. No. 6e-08;
  Matches   73;  Conservative   43;  Mismatches  108;  Indels   54;  Gaps   14;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWL 58
           : | :  :||| :|    |::: : |  :   :      | | ::    |       :||
Db      12 LCLFILAILLCSLA--LGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPR------VEWK 63

Qy      59 LSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTY 118
             |  :    :: |: :||   |    : || |    | :|    :||    | |||
Db      64 FDQGDTTR----LVCYN-NKITASY----EDRVTF----LPTGITFKSVTR---EDTGTY 107

Qy     119 QCKVKKAPG--VGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQ 176
           | | :  |   |  |::| ||: ||     :  |  |||   | |  ::|| |  | |
Db     108 TCMVSEEGGNSYGEVKVKLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWF 167

Qy     177 K--------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
           |                |||  :      |:    :|  :     :| |||  :|  |:
Db     168 KDGIVMPTNPKSTRAFSNSSYVLNPTTGELVFDPLSASD-----TGEYSCEARNGYGTPM 222

Qy     223 C--LLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVF 258
             :|:: |      | |  |   ||: |:|:|::||
Db     223 TSNAVRMEAV--ERNVGVIVAA---VLVTLILLGILVF 255


RESULT   5
JAM2_HUMAN
ID   JAM2_HUMAN     STANDARD;      PRT;   298 AA.
AC   P57087;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Junctional adhesion molecule 2 precursor (Vascular endothelial
DE   junction-associated molecule) (VE-JAM).
GN   JAM2 OR C21ORF43.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Vascular endothelial cells;
RX   MEDLINE=20317114; PubMed=10779521;
RA   Palmeri D., van Zante A., Huang C.C., Hemmerich S., Rosen S.D.;
RT   "Vascular endothelial junction-associated molecule, a novel member of
RT   the immunoglobulin superfamily, is localized to intercellular
RT   boundaries of endothelial cells.";
RL   J. Biol. Chem. 275:19139-19145(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20507930; PubMed=10945976;
RA   Cunningham S.A., Arrate M.P., Rodriguez J.M., Bjercke R.J.,
RA   Vanderslice P., Morris A.P., Brock T.A.;
RT   "A novel protein with homology to the junctional adhesion molecule:
RT   Characterization of leukocyte interactions.";
RL   J. Biol. Chem. 275:34750-34756(2000).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE PROCESSES OF LYMPHOCYTE HOMING TO
CC       SECONDARY LYMPHOID ORGANS.
```

```
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: PROMINENTLY EXPRESSED ON HIGH ENDOTHELIAL
CC       VENULES BUT IS ALSO PRESENT ON THE ENDOTHELIA OF OTHER VESSELS.
CC       LOCALIZED TO THE INTERCELLULAR BOUNDARIES OF HIGH ENDOTHELIAL
CC       CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF255910; AAF81223.1; -.
DR   EMBL; AY016009; AAG49022.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Signal.
FT   SIGNAL       1    20       POTENTIAL.
FT   CHAIN       21   298       JUNCTIONAL ADHESION MOLECULE 2.
FT   DOMAIN      21   238       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   239   259       POTENTIAL.
FT   DOMAIN     260   298       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      43   116       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN     148   221       IG-LIKE C2-TYPE DOMAIN.
FT   DISULFID    50   109       POTENTIAL.
FT   DISULFID   155   214       POTENTIAL.
FT   CARBOHYD    98    98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   187   187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   236   236       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   298 AA;  33207 MW;  CA78E518E22DCAEE CRC64;


  Query Match             9.8%;  Score 186;  DB 1;  Length 298;
  Best Local Similarity   24.5%;  Pred. No. 4.1e-07;
  Matches   82;  Conservative   49;  Mismatches  136;  Indels   68;  Gaps   12;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           : ||| ::::           |    :|::   : : | | |:    |:      :||
Db      10 LLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACK---TPKKTVSSRLEW--- 63

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQ-----DLKGRVHFTSNDLKSGDASINVTNLQLSDI 115
               :|:        |  :   |||     | | |           | :| : |:  ||
Db      64 ----KKL--------GRSVSFVYYQQTLQGDFKNRAEMI-------DFNIRIKNVTRSDA 104

Qy     116 GTYQCKVKKAPGVGNKKIQ-----LTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLP 170
           | |:|:|  ||    : ::     | ||: |:   | |  |   |   :|:|: |||:
Db     105 GKYRCEV-SAPSEQGQNLEEDTVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPA 163

Qy     171 LLYEWQK-----LSN----SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSD 221
             | | |     | |    ||   : :     :  :     |   :| |||  :| ||
Db     164 PEYTWFKDGIRLLENPRLGSQSTNSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYR 223

Qy     222 QCLLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDV 281
           :|  :   |   | :| || ||: | | : : || |  |: :|
Db     224 RCPGKRMQVDDLNISGIIA-AVVVVALVISVCGLGV--CYAQR---------------- 263

Qy     282 PPPKSRTSTARSYLGSNHSSLGSMSPSNMEGYSKT 316
               |   |:  || ||  :    |   ::|:
Db     264 ---KGYFSKETSFQKSNSSSKATTMSENDFKHTKS 295


RESULT   6
JAM1_BOVIN
ID   JAM1_BOVIN     STANDARD;      PRT;   298 AA.
AC   Q9XT56;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Junctional adhesion molecule 1 precursor (JAM).
GN   JAM1.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99323940; PubMed=10395639;
RA   Ozaki H., Ishii K., Horiuchi H., Arai H., Kawamoto T., Okawa K.,
RA   Iwamatsu A., Kita T.;
RT   "Cutting edge: combined treatment of TNF-alpha and IFN-gamma causes
RT   redistribution of junctional adhesion molecule in human endothelial
RT   cells.";
RL   J. Immunol. 163:553-557(1999).
CC   -!- FUNCTION: PLAYS A ROLE IN REGULATING MONOCYTE TRANSMIGRATION
CC       INVOLVED IN INTEGRITY OF EPITHELIAL BARRIER. INVOLVED IN PLATELET
CC       ACTIVATION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: LOCALIZED AT TIGHT JUNCTIONS OF BOTH
CC       EPITHELIAL AND ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF111714; AAD42051.1; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 1.
KW   Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW   Repeat; Signal.
FT   SIGNAL       1    24       POTENTIAL.
FT   CHAIN       25   298       JUNCTIONAL ADHESION MOLECULE 1.
FT   DOMAIN      25   237       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   238   258       POTENTIAL.
FT   DOMAIN     259   298       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      42   115       IG-LIKE V-TYPE DOMAIN 1.
FT   DOMAIN     145   218       IG-LIKE V-TYPE DOMAIN 2.
FT   DISULFID    49   108       POTENTIAL.
FT   DISULFID   152   211       POTENTIAL.
FT   CARBOHYD   184   184       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   298 AA;  32456 MW;  714FE1C1714769A2 CRC64;


  Query Match             9.0%;  Score 170;  DB 1;  Length 298;
  Best Local Similarity   25.5%;  Pred. No. 6.7e-06;
  Matches   70;  Conservative   41;  Mismatches  113;  Indels   50;  Gaps   13;

Qy       3 LLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWLLS 60
             ||   ::|| :| | |  :: | | ::   :   | | | ::    |       :||  :
Db      13 LLFTSMILCSLA-LGRG-AVQTYEPVVRVPENNPAKLSCSYSGFSSPR------VEWKFT 64

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
               | :     :: |: :||   |    : || |:       | |   ::   | | | |
Db      65 HGDIRG----LVCYN-NKITASY----ENRVTFS-------DTGITFHSVTRKDTGMYTC 108

Qy     121 KVKKAPG--VGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQK- 177
               |    |   |   :|| ||: ||     |  |  ||    | |  ::|| |  |:| |
```

```
Db    109 MVSDEGGNTYGEVTVQLIVLVPPSKPTINVPSSVTIGTRAVLTCSERDGSPPSEYKWFKD 168

Qy    178 ------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCL 224
                       |||    :  | : |  :|     |   :| ::|  :|   |
Db    169 GVEMPLEPKSNRAFSNS----SYTLNQKTGELI-FDPVSASDTGDFTCQAQNGYASP--- 220

Qy    225 LRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVF 258
          :: | |        : | |  | : |:|:| ::|
Db    221 VKSDTVHMDAVELNVGGIVAAVFVTLILLGALIF 254


RESULT   7
CD80_MOUSE
ID   CD80_MOUSE     STANDARD;      PRT;   306 AA.
AC   Q00609;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   T lymphocyte activation antigen CD80 precursor (Activation B7-1
     antigen) (B7).
GN   CD80 OR B7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-cell;
RX   MEDLINE=91341422; PubMed=1714935;
RA   Gray G.S., Freeman G.J., Gimmi C.D., Lombard D.B., Zhou L.J.,
RA   White M., Fingeroth J.D., Gribben J.G., Nadler L.M.;
RT   "Structure, expression, and T cell costimulatory activity of the
RT   murine homologue of the human B lymphocyte activation antigen B7.";
RL   J. Exp. Med. 174:625-631(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=B-cell;
RX   MEDLINE=93307789; PubMed=7686531;
RA   Selvakumar A., White P.C., Dupont B.;
RT   "Genomic organization of the mouse B-lymphocyte activation antigen
RT   B7.";
RL   Immunogenetics 38:292-295(1993).
CC   -!- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T
CC       LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE
CC       PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS
CC       RECEPTOR.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED ON ACTIVATED B CELLS, GAMMA
CC       INTERFERON STIMULATED MONOCYTES AND NONCIRCULATING B-CELL
CC       MALIGNANCIES.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED BETWEEN 4 AND 12 HOURS POST-
CC       ACTIVATION. PROTEIN WAS DETECTED AT CELL SURFACE AT 24 HOURS AND
CC       IT'S EXPRESSION WAS MAXIMAL FROM 48 TO 72 HOURS POST-ACTIVATION.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X60958; CAA43291.1; -.
DR   EMBL; L12589; AAA37240.1; ALT_SEQ.
DR   EMBL; L12585; AAA37240.1; JOINED.
DR   EMBL; L12586; AAA37240.1; JOINED.
DR   EMBL; L12587; AAA37240.1; JOINED.
DR   EMBL; L12588; AAA37240.1; JOINED.
DR   PIR; S17291; S17291.
DR   MGD; MGI:101775; Cd80.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane;
KW   Receptor.
FT   SIGNAL        1     37
FT   CHAIN        38    306       T LYMPHOCYTE ACTIVATION ANTIGEN CD80.
FT   DOMAIN       38    246       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    247    268       POTENTIAL.
FT   DOMAIN      269    306       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       47    126       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      158    226       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      227    246       IG-HINGE LIKE (POTENTIAL).
FT   DISULFID     54    119       POTENTIAL.
FT   DISULFID    165    219       POTENTIAL.
FT   CARBOHYD     93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     99     99       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    149    149       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    210    210       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    214    214       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   306 AA;  34589 MW;  1DBADE0931B84C62 CRC64;


  Query Match             8.5%;  Score 161.5;  DB 1;  Length 306;
  Best Local Similarity   23.2%;  Pred. No. 3.1e-05;
  Matches   68;  Conservative   56;  Mismatches  110;  Indels   59;  Gaps   14;

Qy      3 LLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPA 62
          |:| |||| :: ::  :     :: : |:  :   ||||:    ||:    | |
Db     21 LILLFVLLIRLSQVSSDV-----DEQLSKSVKDKVLLPCRYNSPHEDESEDRIYW----- 70

Qy     63 DNQKVDQVII-LYSGD-KIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
            || |:|:: : :|  |:: :|    | |  : :        |: :  | ||| ||| |
Db     71 --QKHDKVVLSVIAGKLKVWPEY----KNRTLYDNTTY-----SLIILGLVLSDRGTYSC 119

Qy    121 KVKKAPGVGNKKIQLTVLLKPS-----GTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEW 175
          |:|    |  :::   |:| |      |    :      :  :: |    |     : |
Db    120 VVQKKER-GTYEVKHLALVKLSIKADFSTPNITESGNPSADTKRITCFASGGFPKPRFSW 178

Qy    176 QKLSNSQKLPTLWLAEMTSPVISVKNAST--EYSGTYSCTVKNRVGSDQCLLRL------ 227
            | | ::|| :       |   :   |:  ::: | : |:|       ||::
Db    179 --LENGRELPGINTTISQDPESELYTISSQLDFNTTRNHTIK-------CLIKYGDAHVS 229

Qy    228 ---------DVVPPSNRAGTIAGAVIGVLLALVLIGLIV--FCCHKK--RREE 267
                  :  | |     : ||  | :: :|:| :|:  || |:   || |
Db    230 EDFTWEKPPEDPPDSKNTLVLFGAGFGAVITVVVIVVIIKCFCKHRSCFRRNE 282


RESULT   8
JAM1_MOUSE
ID   JAM1_MOUSE     STANDARD;      PRT;   300 AA.
AC   O88792;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Junctional adhesion molecule 1 precursor (JAM).
GN   JAM1 OR JCAM1 OR JCAM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98327120; PubMed=9660867;
RA   Martin-Padura I., Lostaglio S., Schneemann M., Williams L., Romano M.,
RA   Fruscella P., Panzeri C., Stoppacciaro A., Ruco L., Villa A.,
RA   Simmons D., Dejana E.;
RT   "Junctional adhesion molecule, a novel member of the immunoglobulin
RT   superfamily that distributes at intercellular junctions and modulates
```

```
RT   monocyte transmigration.";
RL   J. Cell Biol. 142:117-127(1998).
CC   -!- FUNCTION: PLAYS A ROLE IN REGULATING MONOCYTE TRANSMIGRATION
CC       INVOLVED IN INTEGRITY OF EPITHELIAL BARRIER. INVOLVED IN PLATELET
CC       ACTIVATION.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: LOCALIZED AT TIGHT JUNCTIONS OF BOTH
CC       EPITHELIAL AND ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U89915; AAC32982.1; -.
DR   MGD; MGI:1321398; Jcam1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00406; IGv; 1.
KW   Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW   Repeat; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    300       JUNCTIONAL ADHESION MOLECULE 1.
FT   DOMAIN       27    238       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    239    259       POTENTIAL.
FT   DOMAIN      260    299       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42    115       IG-LIKE V-TYPE DOMAIN 1.
FT   DOMAIN      145    219       IG-LIKE V-TYPE DOMAIN 2.
FT   DISULFID     49    108       POTENTIAL.
FT   DISULFID    152    212       POTENTIAL.
FT   CARBOHYD     42     42       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   300 AA;  32368 MW;  391F3E48FF3B97EC CRC64;


  Query Match             8.5%;  Score 160.5;  DB 1;  Length 300;
  Best Local Similarity   25.1%;  Pred. No. 3.6e-05;
  Matches   64;  Conservative   38;  Mismatches  108;  Indels   45;  Gaps   11;

Qy      21 SITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWLLSPADNQKVDQVIILYSGDK 78
           |: | :  ::  : |:  | | ::    |       :||       |     :: |:  :
Db      29 SVYTAQSDVQVPENESIKLTCTYSGFSSPR------VEWKFV----QGSTTALVCYN-SQ 77

Qy      79 IYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPG--VGNKKIQLT 136
           |   |      || |:|       : |  :::   | | | | | :  |   |   | ||
Db      78 ITAPY----ADRVTFSS-------SGITFSSVTRKDNGEYTCMVSEEGGQNYGEVSIHLT 126

Qy     137 VLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQK------LSNSQKLPTLWLA 190
           ||: ||     |  |  |||   | |   :|| |  | | |       ::::|      :
Db     127 VLVPPSKPTISVPSSVTIGNRAVLTCSEHDGSPPSEYSWFKDGISMLTADAKKTRAFMNS 186

Qy     191 EMTSPVISVKNASTEY-------SGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIAGAV 243
             |   |  |:     :       || | |  :|  |:    :| :          : | |
Db     187 SFT---IDPKSGDLIFDPVTAFDSGEYYCQAQNGYGT---AMRSEAAHMDAVELNVGGIV 240

Qy     244 IGVLLALVLIGLIVF 258
             ||: |:|:||::|
Db     241 AAVLVTLILLGLLIF 255


RESULT   9
PVR2_HUMAN
ID   PVR2_HUMAN     STANDARD;      PRT;   538 AA.
AC   Q92692; O75455;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Poliovirus receptor related protein 2 precursor (Herpes virus entry
DE   mediator B) (HveB) (Nectin 2) (CD112 antigen).
GN   PVRL2 OR PRR2 OR HVEB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM DELTA).
RX   MEDLINE=95347610; PubMed=7622062;
RA   Eberle F., Dubreuil P., Mattei M.-G., Devilard E., Lopez M.;
RT   "The human PRR2 gene, related to the human poliovirus receptor gene
RT   (PVR), is the true homolog of the murine MPH gene.";
RL   Gene 159:267-272(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM ALPHA).
RX   MEDLINE=98321161; PubMed=9657005;
RA   Warner M.S., Geraghty R.J., Martinez W.M., Montgomery R.I.,
RA   Whitbeck J.C., Xu R., Eisenberg R.J., Cohen G.H., Spear P.G.;
RT   "A cell surface protein with herpesvirus entry activity (HveB) confers
RT   susceptibility to infection by mutants of herpes simplex virus type
RT   1, herpes simplex virus type 2, and pseudorabies virus.";
RL   Virology 246:179-189(1998).
RN   [3]
RP   SEQUENCE OF 31-538 FROM N.A.
RA   Yoshiura K., Murray J.C.;
RT   "A transcriptional map in the region of 19q13 derived using direct
RT   sequencing and exon trapping.";
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 449-538 FROM N.A.
RX   MEDLINE=99449047; PubMed=10520737;
RA   Freitas E.M., Zhang W.J., Lalonde J.P., Tay G.K., Gaudieri S.,
RA   Ashworth L.K., Van Bockxmeer F.M., Dawkins R.L.;
RT   "Sequencing of 42kb of the APO E-C2 gene cluster reveals a new gene:
RT   PEREC1.";
RL   DNA Seq. 9:89-101(1998).
CC   -!- FUNCTION: RECEPTOR FOR ALPHAHERPESVIRUS (HSV-1, HSV-2 AND
CC       PSEUDORABIES VIRUS) ENTRY INTO CELLS.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA AND DELTA (SHOWN HERE);
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD112 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/204270028_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X80038; CAA56342.1; -.
DR   EMBL; AF058448; AAC23797.1; -.
DR   EMBL; AF044968; AAC82348.1; -.
DR   EMBL; AF044962; AAC82348.1; JOINED.
DR   EMBL; AF044963; AAC82348.1; JOINED.
DR   EMBL; AF044964; AAC82348.1; JOINED.
DR   EMBL; AF044966; AAC82348.1; JOINED.
DR   EMBL; AF044967; AAC82348.1; JOINED.
DR   EMBL; AF050154; AAD02503.1; -.
DR   MIM; 600798; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00409; IG; 1.
```

```
KW   Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;
KW   Repeat; Alternative splicing.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32    538       POLIOVIRUS RECEPTOR RELATED PROTEIN 2.
FT   DOMAIN       32    360       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    361    381       POTENTIAL.
FT   DOMAIN      382    538       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       47    147       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      176    245       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      276    336       IG-LIKE C2-TYPE DOMAIN 2.
FT   DISULFID     54    140       BY SIMILARITY.
FT   DISULFID    183    238       BY SIMILARITY.
FT   DISULFID    283    329       BY SIMILARITY.
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    324    324       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    351    479       NTAGAGATGGIIGGIIAAIIATATVAATGILICRQQRKEQTL
FT                                QGAEEDEDLEGPPSYKPPTPKAKLEAQEMPSQLFTLGASEH
FT                                SPLKTPYFDAGASCTEQEMPRYHELPTLEERSGPLHPGATS
FT                                LGSPIP -> RASPRDVGPLVWGAVGGTLLVLLLLAGGSLA
FT                                FILLRVRRRRKSPGGAGGGASGDGGFYDPKAQVLGNGDPVF
FT                                WTPVVPGPMEPDGKDEEEEEEEEKAEKGLMLPPPPALEDDM
FT                                ESQLDGSLISRRAVYV (IN ISOFORM ALPHA).
FT   VARSPLIC    480    538       MISSING (IN ISOFORM ALPHA).
SQ   SEQUENCE   538 AA;  57742 MW;  3AE4F83E92F6F624 CRC64;


  Query Match             8.0%;  Score 152.5;  DB 1;  Length 538;
  Best Local Similarity   23.7%;  Pred. No. 0.00031;
  Matches   86;  Conservative   51;  Mismatches  155;  Indels   71;  Gaps   17;

Qy      27 QMIEKAKGETAYLPCRFTLGP---------EDQGPLDIEWLLSPADNQKVDQVIILYSGD 77
           ::| | | :       |: |          | : |  | || |     |  ||    ||
Db     157 RVIAKPKNQAEAQKVTFSQDPTTVALCISKEGRPPARISWLSSLDWEAKETQV----SG- 211

Qy      78 KIYDDYYQDLKGRVHFTSND--LKSGDASINVTNLQLSDIGTYQCKVK----KAPGVGNK 131
                | | |  ||    : || |          |  |  |||:    : | :
Db     212 --------TLAGTVTVTSRFTLVPSGRA----------DGVTVTCKVEHESFEEPAL--I 251

Qy     132 KIQLTVLLKPSGTRCYVDGSEEIG-NDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLA 190
           : |:|   |  :    | :  :|  |  | |: :    |  |:|   | :   ||  :|
Db     252 PVTLSVRYPPEVSISGYDDNWYLGRTDATLSCDVRSNPEPTGYDWSTTSGT--FPTSAVA 309

Qy     191 EMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGT-----IAGAVIG 245
           : :  ||      : :: |: ||| | ||  :    : |    | ||      | | :|
Db     310 QGSQLVIHA--VDSLFNTTFVCTVTNAVGMGRAEQVIFVRETPNTAGAGATGGIIGGIIA 367

Qy     246 VLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVP---PPKSRTSTARSYLGSNHSSL 302
           ::|  :    :  | ::|:|:  :     :  |  |   ||  :       : |   :|
Db     368 AIIATAVAATGILICRQQRKEQTLQGAEEDEDLEGPPSYKPPTPKAKLEAQEMPSQLFTL 427

Qy     303 GSMSPSNMEGYSKTQY----NQVPSEDFERAPQSPTL-----PLAKVAAPNLSRMGAVPV 353
           |:   | :    || |         ::  |  : |||     ||     |  : :|: |:
Db     428 GASEHSPL----KTPYFDAGASCTEQEMPRYHELPTLEERSGPL----HPGATSLGS-PI 478

Qy     354 MIP 356
           :|
Db     479 PVP 481


RESULT  10
CEA1_HUMAN
ID   CEA1_HUMAN     STANDARD;      PRT;   526 AA.
AC   P13688;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Carcinoembryonic antigen-related cell adhesion molecule 1 precursor
DE   (Biliary glycoprotein 1) (BGP-1) (Antigen CD66) (CD66a antigen).
GN   CEACAM1 OR BGP OR BGP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=88320555; PubMed=2457922;
RA   Hinoda Y., Neumaier M., Hefta S.A., Drzeniek Z., Wagener C.,
RA   Shively L., Hefta L.J.F., Shively J.E., Paxton R.J.;
RT   "Molecular cloning of a cDNA coding biliary glycoprotein I: primary
RT   structure of a glycoprotein immunologically crossreactive with
RT   carcinoembryonic antigen.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:6959-6963(1988).
RN   [2]
RP   REVISION.
RA   Hinoda Y., Neumaier M., Hefta S.A., Drzeniek Z., Wagener C.,
RA   Shively L., Hefta L.J.F., Shively J.E., Paxton R.J.;
RL   Proc. Natl. Acad. Sci. U.S.A. 86:1668-1668(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89139550; PubMed=2537311;
RA   Barnett T.R., Kretschmer A., Austen D.A., Goebel S.J., Hart J.T.,
RA   Elting J.J., Kamarck M.E.;
RT   "Carcinoembryonic antigens: alternative splicing accounts for the
RT   multiple mRNAs that code for novel members of the carcinoembryonic
RT   antigen family.";
RL   J. Cell Biol. 108:267-276(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=91222218; PubMed=2025273;
RA   Kuroki M., Arakawa F., Matsuo Y., Oikawa S., Nakazato H.,
RA   Matsuoka Y.;
RT   "Three novel molecular forms of biliary glycoprotein deduced from
RT   cDNA clones from a human leukocyte library.";
RL   Biochem. Biophys. Res. Commun. 176:578-585(1991).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA   Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA   Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA   Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (ISOFORM A) AND
CC       SECRETED (ISOFORMS G, H AND I).
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; A (shown here), G, H and I; are
CC       produced by alternative splicing.
CC   -!- DISEASE: INCREASED SERUM LEVELS OF BGP-1 ARE FOUND IN INDIVIDUALS
CC       SUFFERING FROM HEPATIC DISORDERS.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
CC       SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD66a entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd66a.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16354; CAA34404.1; -.
DR   EMBL; J03858; AAA51826.1; -.
DR   EMBL; D90311; BAA14341.1; -.
DR   EMBL; D90312; BAA14342.1; -.
DR   EMBL; D90313; BAA14343.1; -.
DR   EMBL; M69176; AAA51825.1; -.
DR   EMBL; M72238; AAA58393.1; -.
DR   EMBL; M72238; AAA58394.1; -.
DR   EMBL; AC004785; AAC18434.1; -.
```

```
DR   EMBL; AC004785; AAC18437.1; -.
DR   EMBL; AC004785; AAC18438.1; -.
DR   EMBL; AC004785; AAC18439.1; -.
DR   PIR; A32164; A32164.
DR   PIR; JH0394; JH0394.
DR   PIR; JH0395; JH0395.
DR   PIR; JH0396; JH0396.
DR   MIM; 109770; -.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00410; IG_like; 1.
DR   SMART; SM00408; IGc2; 2.
KW   Immunoglobulin domain; Glycoprotein; Transmembrane; Signal;
KW   Repeat; Alternative splicing.
FT   SIGNAL        1     34
FT   CHAIN        35    526       CARCINOEMBRYONIC ANTIGEN-RELATED CELL
FT                                 ADHESION MOLECULE 1.
FT   DOMAIN       35    428       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    429    452       POTENTIAL.
FT   DOMAIN      453    526       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       35    142       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      160    222       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      252    306       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN      341    403       IG-LIKE C2-TYPE DOMAIN 3.
FT   DISULFID    167    215       PROBABLE.
FT   DISULFID    259    299       PROBABLE.
FT   DISULFID    348    396       PROBABLE.
FT   MOD_RES      35     35       PYRROLIDONE CARBOXYLIC ACID.
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    111    111       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    115    115       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    152    152       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    182    182       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    208    208       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    254    254       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    302    302       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    351    351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    363    363       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    378    378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    320    321       EL -> GK (IN ISOFORM H).
FT   VARSPLIC    322    526       MISSING (IN ISOFORM H).
FT   VARSPLIC    321    351       LSPVVAKPQIKASKTTVTGDKDSVNLTCSTN -> SPVLGE
FT                                 DEAVPGQHHPQHKPCQEGGCWDVLV (IN ISOFORM I).
FT   VARSPLIC    352    526       MISSING (IN ISOFORM I).
FT   VARSPLIC    416    417       YN -> CK (IN ISOFORM G).
FT   VARSPLIC    418    526       MISSING (IN ISOFORM G).
SQ   SEQUENCE   526 AA;  57560 MW;  CAD1B2328D069AF8 CRC64;


  Query Match             8.0%;  Score 151;  DB 1;  Length 526;
  Best Local Similarity   22.4%;  Pred. No. 0.00039;
  Matches   76;  Conservative   56;  Mismatches  133;  Indels   74;  Gaps   18;

Qy      49 DQGPLDIEWLLSPADNQKVDQVI--ILYSGD----KIYDDYYQ---DLKGRVHFTSN--- 96
           | || : | : :|    : | |   : |  |       | ||:   :|    :  ||
Db     209 DTGPYECE-IQNPVSANRSDPVTLNVTYGPDTPTISPSDTYYRPGANLSLSCYAASNPPA 267

Qy      97 --------DLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLL--------K 140
                    :    : : |: ::: |:| |    :    |:    |:::       |
Db     268 QYSWLINGTFQQSTQELFIPNITVNNSGSYTCHANNSVTGCNRTTVKTIIVTELSPVVAK 327

Qy     141 P--SGTRCYVDGSEEIGNDFKLKCEPKEGSLPL--LYEWQKLSNSQKLPTLWLAEMTSPV 196
           |    ::  | | ::  |   | |   :  : :   :: | | :|:::        :
Db     328 PQIKASKTTVTGDKDSVN---LTCSTNDTGISIRWFFKNQSLPSSERMK----LSQGNTT 380

Qy     197 ISVKNASTEYSGTYSCTVKNRVG---SDQCLLRL--DVVPPSN--RAGTIAGAVIGV--L 247
           :|:     | :||| | | | :    ||  :| :  : :|  |    | ||| ||||  |
Db     381 LSINPVKREDAGTYWCEVFNPISKNQSDPIMLNVNYNALPQENGLSPGAIAGIVIGVVAL 440

Qy     248 LALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHSSLGSMSP 307
           :||: : |  |     |      ::    |: |  |                |||:   |  |
Db     441 VALIAVALACFLHFGKTGRASDQR----DLTEHKPSV-----------SNHTQDHSNDP 484

Qy     308 SNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLS 346
            |       : |:|         | || | :  |:|:|:
Db     485 PN-------KMNEVTYSTLNFEAQQPTQPTS--ASPSLT 514


RESULT  11
PEC1_PIG
ID   PEC1_PIG       STANDARD;      PRT;   740 AA.
AC   Q95242;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Platelet endothelial cell adhesion molecule precursor (PECAM-1)
DE   (CD31 antigen).
GN   PECAM1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nasu K.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE EXPRESSED ON
CC       PLATELETS AND AT ENDOTHELIAL CELL INTERCELLULAR JUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X98505; CAA67129.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
KW   Repeat; Signal.
FT   SIGNAL        1     27       BY SIMILARITY.
FT   CHAIN        28    740       PLATELET ENDOTHELIAL CELL ADHESION
FT                                 MOLECULE.
FT   DOMAIN       28    602       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    603    621       POTENTIAL.
FT   DOMAIN      622    740       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       50    116       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      145    213       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN      249    311       IG-LIKE C2-TYPE DOMAIN 3.
FT   DOMAIN      340    394       IG-LIKE C2-TYPE DOMAIN 4.
FT   DOMAIN      425    484       IG-LIKE C2-TYPE DOMAIN 5.
FT   DOMAIN      517    580       IG-LIKE C2-TYPE DOMAIN 6.
FT   CARBOHYD     52     52       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     84     84       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    301    301       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    357    357       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    372    372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    436    436       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    456    456       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    552    552       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   740 AA;  82378 MW;  F312DC62C4B4A217 CRC64;


  Query Match             7.9%;  Score 149;  DB 1;  Length 740;
  Best Local Similarity   23.5%;  Pred. No. 0.00087;
  Matches   58;  Conservative   44;  Mismatches  115;  Indels   30;  Gaps    8;

Qy      84 YQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC---KVKKAPGVGNKKIQLTVLLK 140
           || |||     | :: | : ::   |    |:  |||         |: :| ::: |:
Db     443 YQLLKGSDLLASQNVSSNEPAVFKDN-PTKDV-EYQCIADNCHSHAGMPSKVLRVKVIAP 500

Qy     141 PSGTRCYVDGSEEI--GNDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVIS 198
             :  : |||:  |    |:|  |||| |: |:: |   ::    : | : |  :
Db     501 VEEVKLSILLSEEVESGQAIVLQCSVKEGSGPITYKFYKEKENKPFHQVTLND-TQAIWH 559

Qy     199 VKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVP-----PSNRAGTIAGAVIGVLLALVLI 253
           || :  | | |   ||    :  |: :::         : | ||  || |::|::|:
Db     560 KPKASKDQEGQYYCLASNRATPSKNFLQSNILAVRVYLAPWKKGLIAVVVIAVIIAVLLL 619

Qy     254 GLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHSSLGSMSPSNMEGY 313
           |   :  | : ::              |     |   | ||:    ::|  | |
Db     620 GARFYFLKKSKAKQM--------------PVEMCRPAAPLLNSNNEK--TLSDPNTEAN 662

Qy     314 SKTQYNQ 320
             ||:
Db     663 RHYGYNE 669


RESULT  12
PEC1_MOUSE
ID   PEC1_MOUSE     STANDARD;      PRT;   727 AA.
AC   Q08481;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Platelet endothelial cell adhesion molecule precursor (PECAM-1)
DE   (CD31 antigen).
GN   PECAM1 OR PECAM-1 OR PECAM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Lung;
RX   MEDLINE=93296179; PubMed=8516303;
RA   Xie Y., Muller W.A.;
RT   "Molecular cloning and adhesive properties of murine
RT   platelet/endothelial cell adhesion molecule 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:5569-5573(1993).
RN   [2]
RP   SEQUENCE OF 18-26.
RC   TISSUE=Heart;
RX   MEDLINE=93035639; PubMed=1415479;
RA   Bogen S.A., Baldwin H.S., Watkins S.C., Albelda S.M., Abbas A.K.;
RT   "Association of murine CD31 with transmigrating lymphocytes following
RT   antigenic stimulation.";
RL   Am. J. Pathol. 141:843-854(1992).
CC   -!- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE EXPRESSED ON
CC       PLATELETS AND AT ENDOTHELIAL CELL INTERCELLULAR JUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L06039; AAA16230.1; -.
DR   MGD; MGI:97537; Pecam.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 3.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00410; IG_like; 2.
KW   Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
KW   Repeat; Phosphorylation; Signal.
FT   SIGNAL       1     17
FT   CHAIN       18    727       PLATELET ENDOTHELIAL CELL ADHESION
FT                               MOLECULE.
FT   DOMAIN       18    590       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    591    609       POTENTIAL.
FT   DOMAIN      610    727       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       40    106       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      135    203       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN      238    300       IG-LIKE C2-TYPE DOMAIN 3.
FT   DOMAIN      329    382       IG-LIKE C2-TYPE DOMAIN 4.
FT   DOMAIN      413    472       IG-LIKE C2-TYPE DOMAIN 5.
FT   DOMAIN      505    568       IG-LIKE C2-TYPE DOMAIN 6.
FT   DISULFID     47     99       POTENTIAL.
FT   DISULFID    142    195       POTENTIAL.
FT   DISULFID    245    293       POTENTIAL.
FT   DISULFID    336    375       POTENTIAL.
FT   DISULFID    420    465       POTENTIAL.
FT   DISULFID    512    561       POTENTIAL.
FT   CARBOHYD     74     74       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    141    141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    309    309       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    345    345       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    540    540       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MOD_RES     702    702       PHOSPHORYLATION (POTENTIAL).
FT   CONFLICT     18     18       E -> T (IN REF. 2).
SQ   SEQUENCE   727 AA;  81262 MW;  34C04752D199BAA5 CRC64;


  Query Match             7.8%;  Score 148.5;  DB 1;  Length 727;
  Best Local Similarity   21.8%;  Pred. No. 0.00093;
  Matches   61;  Conservative   51;  Mismatches  127;  Indels   41;  Gaps   10;

Qy      84 YQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKA---PGVGNKKIQLTVLLK 140
           |  :| :  | : :: | | :   |:    |: |||:       | | :: ::: |:
Db     431 YHLMKAKSDFQTLEVTSNDPA-TFTDKPTRDM-EYQCRADNCHSHPAVFSEILRVRVIAP 488

Qy     141 PSGTRCYVDGSEEI--GNDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVIS 198
             :  | |:  |::  |:|  |||: |: ::: |    :      :   |
Db     489 VDEVVISILSSNEVQSGSEMVLRCSVKEGTSPITFQFYKEKEDRPFHQA-VVNDTQAFWH 547

Qy     199 VKNASTEYSGTYSCTVKNRVG-------SDQCLLRLDVVPPSNRAGTIAGAVIGVLLALV 251
           | || :  | | ||  ||         |    :|: : |   : | ||  ||||::| :
Db     548 NKQASKKQEGQYYCTASNRASSMRTSPRSSTLAVRVFLAP--WKKGLIAVVVIGVVIATL 605

Qy     252 LIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHSSLGSMSPSNME 311
           ::    :  | : ::|               |   :  |   | ||   :   | ::|
Db     606 IVAAKCYFLRKAKAKQK--------------PVEMSRPAAPLLNSNSEKI---SEPSVE 647

Qy     312 GYSKTQYNQVPSEDF------ERAPQSPTLPLAKVAAPNL 345
             |   |: |   |        : ||:  :   :|   :|
Db     648 ANSHYGYDDVSGNDAVKPINQNKDPQNMDVEYTEVEVSSL 687


RESULT  13
ICCR_DROME
ID   ICCR_DROME     STANDARD;      PRT;   764 AA.
```

```
AC   Q08180;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Irregular chiasm C-roughest protein precursor (IRREC protein).
GN   RST.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94102535; PubMed=7503814;
RA   Ramos R.G., Igloi G.L., Lichte B., Baumann U., Maier D.,
RA   Schneider T., Brandstaetter J.H., Froehlich A., Fischbach K.-F.;
RT   "The irregular chiasm C-roughest locus of Drosophila, which affects
RT   axonal projections and programmed cell death, encodes a novel
RT   immunoglobulin-like protein.";
RL   Genes Dev. 7:2533-2547(1993).
CC   -!- FUNCTION: REQUIRED FOR CORRECT AXONAL PATHWAY FORMATION IN
CC       THE OPTIC LOBE AND FOR PROGRAMMED CELL DEATH IN THE DEVELOPING
CC       RETINA.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: POSTEMBRYONIC EXPRESSION IS STRONG IN THE
CC       DEVELOPING OPTIC LOBE AND IN THE EYE IMAGINAL DISC.
CC   -!- DEVELOPMENTAL STAGE: STRONGLY EXPRESSED IN EMBRYOS. ALSO FOUND
CC       IN LATE LARVAL AND PUPAL STAGES.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC   -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z21641; CAA79756.1; -.
DR   EMBL; L11040; AAA16632.1; -.
DR   PIR; A49448; A49448.
DR   PIR; S34129; S34129.
DR   HSSP; P56276; 1TLK.
DR   FlyBase; FBgn0003285; rst.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00410; IG_like; 4.
DR   SMART; SM00408; IGc2; 1.
KW   Transmembrane; Immunoglobulin domain; Glycoprotein; Signal; Repeat;
KW   Cell adhesion.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    764       IRREGULAR CHIASM C-ROUGHEST PROTEIN.
FT   DOMAIN       20    533       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    534    556       POTENTIAL.
FT   DOMAIN      557    764       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       32    115       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      137    221       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN      245    261       GLY-RICH.
FT   DOMAIN      265    332       IG-LIKE C2-TYPE DOMAIN 3.
FT   DOMAIN      353    414       IG-LIKE C2-TYPE DOMAIN 4.
FT   DOMAIN      432    515       IG-LIKE C2-TYPE DOMAIN 5.
FT   DOMAIN      637    660       GLN-RICH (OPA-REPEAT).
FT   CARBOHYD    211    211       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    393    393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    507    507       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   764 AA;  82947 MW;  262225D2B2A1C181 CRC64;
```

```
  Query Match             7.8%;  Score 148.5;  DB 1;  Length 764;
  Best Local Similarity   18.0%;  Pred. No. 0.00099;
  Matches   71;  Conservative   62;  Mismatches  122;  Indels  139;  Gaps   14;

Qy      15 DLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQVIILY 74
           |:: : |     | :|   |    | |     |:     :| |:  |:|
Db     341 DISYAPSFRQRPQSMEADVGSVVSLTCEVDSNPQP----EIVWIQHPSD---------- 385

Qy      75 SGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCK------------ 121
                      ||  || :|     : :|:|      | | ||
Db     386 --------------RVVGTSTNL-----TFSVSN---ETAGRYYCKANVPGYAEISADA 422

Qy     122 ---VKKAPGVGNKKIQLTVLLKPSGTRCYVDG---------------SEEIGNDFKLKC 162
             :| :| :|::: |  ::   :   |:                   | | |:|: :
Db     423 YVYLKGSPAIGSQRTQYGLVGDTARIECFASSVPRARHVSWTFNGQEISSESGHDYSILV 482

Qy     163 EPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
           :  | :                          : :::   : | |:||| |  |:|
Db     483 DAVPGGV----------------------KSTLIIRDSQAYHYGKYNCTVVNDYGNDV 518

Qy     223 CLLRLDVVPPSNRAGTIAG--AVIGVLLALVLIGLIVFCCHKKRR--------EEKYEK- 271
             ::|      :   || |  :|:  || | :: ::   | |: :        | :  |
Db     519 AEIQLQAKKSVSLLMTIVGGISVVAFLLVLTILVVVYIKCKKRTKLPPADVISEHQITKN 578

Qy     272 ------------EVHHDIREDV--------------PPPKSRTSTARSYLGSNHSSLG 303
                       : |:: |:               |||:  |: :     || ||
Db     579 GGVSCKLEPGDRTSNYSDLKVDISGGYVPYGDYSTHYSPPPQYLTTCSTK---SNGSS-- 633

Qy     304 SMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPL 337
           ::  :| :   : |  |  |   :   |: |||:
Db     634 TIMQNNHQNQLQLQQQQQQSH-HQHHTQTTTLPM 666
```

```
RESULT  14
BUTY_BOVIN
ID   BUTY_BOVIN     STANDARD;      PRT;   526 AA.
AC   P18892; O18955; O18959;
DT   01-NOV-1990 (Rel. 16, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Butyrophilin precursor (BT).
GN   BTN1A1 OR BTN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 27-47.
RX   MEDLINE=90354441; PubMed=2387867;
RA   Jack L.J.W., Mather I.H.;
RT   "Cloning and analysis of cDNA encoding bovine butyrophilin, an apical
RT   glycoprotein expressed in mammary tissue and secreted in association
RT   with the milk-fat globule membrane during lactation.";
RL   J. Biol. Chem. 265:14481-14486(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HOLSTEIN-FRIESIAN;
RA   Davey H.W., Ogg S.L., Husaini Y., Snell R.G., Korobko I.V.,
RA   Mather I.H., Wilkins R.J.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Seyfert H., Luethen F.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   CARBOHYDRATE-LINKAGE SITES.
RX   MEDLINE=95293916; PubMed=7775382;
RA   Sato T., Takio K., Kobata A., Greenwalt D.E., Furukawa K.;
RT   "Site-specific glycosylation of bovine butyrophilin.";
RL   J. Biochem. 117:147-157(1995).
CC   -!- FUNCTION: MAY FUNCTION IN THE SECRETION OF MILK-FAT DROPLETS. IT
```

```
CC       MAY ACT AS A SPECIFIC MEMBRANE-ASSOCIATED RECEPTOR FOR THE
CC       ASSOCIATION OF CYTOPLASMIC DROPLETS WITH THE APICAL PLASMA
CC       MEMBRANE.
CC   -!- SUBUNIT: SEEMS TO ASSOCIATE WITH XANTHINE DEHYDROGENASE/OXIDASE
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MAMMARY TISSUE AND SECRETED IN
CC       ASSOCIATION WITH THE MILK-FAT-GLOBULE MEMBRANE DURING LACTATION.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. BTN/MOG
CC       SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35551; AAB39766.1; -.
DR   EMBL; AF005497; AAB62889.1; -.
DR   EMBL; Z93323; CAB07533.1; -.
DR   PIR; A37821; A37821.
DR   InterPro; IPR001870; Gamma_carbxylse.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   InterPro; IPR003877; SPRY.
DR   InterPro; IPR003878; SPRY_domain.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF00622; SPRY; 1.
DR   SMART; SM00406; IGv; 1.
DR   SMART; SM00449; SPRY; 1.
KW   Transmembrane; Glycoprotein; Immunoglobulin domain; Signal.
FT   SIGNAL        1     26
FT   CHAIN        27    526       BUTYROPHILIN.
FT   DOMAIN       27    242       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    243    269       POTENTIAL.
FT   DOMAIN      270    526       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     55     55       N-LINKED (GLCNAC. . .) (COMPLEX).
FT   CARBOHYD    215    215       N-LINKED (GLCNAC. . .) (HYBRID).
FT   CONFLICT     35     35       Q -> P (IN REF. 3).
FT   CONFLICT    230    230       E -> D (IN REF. 1).
SQ   SEQUENCE   526 AA;  59276 MW;  A14126802BD19284 CRC64;


  Query Match             7.8%;  Score 148;  DB 1;  Length 526;
  Best Local Similarity   22.5%;  Pred. No. 0.00066;
  Matches   67;  Conservative   53;  Mismatches  126;  Indels   52;  Gaps   11;

Qy       4 LLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWL---LS 60
           || |:||      :    :  |:: |    || | |||| :     :| ::: |    :|
Db      13 LLIFILLQLPKLDSAPFDVIGPQEPILAVVGEDAELPCRLSPNVSAKG-MELRWFREKVS 71

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||       | :   | :   :   : :|||    : :  |  :: :  :: || | |:|
Db      72 PA-------VFVSREGQEQEGEEMAEYRGRVSLVEDHIAEGSVAVRIQEVKASDDGEYRC 124

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKC-------EPK------EG 167
           ::       : | |     |:  ::    :   : :|:|       ||:      |
Db     125 FFRQDENYEEAIVHLKV--AALGSDPHISMKVQESGEIQLECTSVGWYPEPQVQWRTHRG 182

Qy     168 SLPLLYEWQKLSNSQKLPTLWLAEMTSPVI----SVKNASTEYSGTYSCTVKNRVGSDQC 223
                 |:  :| |:      |  : : ||    |:||         || ::| :   :
Db     183 E-----EFPSMSESRNPDEEGLFTVRASVIIRDSSMKNV--------SCCIRNLLLGQEK 229

Qy     224 LLRLDVVP---PSNRAGTIAGAVIGVLLALVLIGLIVFCCH------KKRREEKYEKE 272
           : : :     |      :| |||| |:| |: || | |          ::|| |   ||
Db     230 EVEVSIPASFFPRLTPWMVAVAVILVVLGLLTIGSIFFTWRLYKERSRQRRNEFSSKE 287


RESULT  15
ECTO_RAT
ID   ECTO_RAT       STANDARD;      PRT;   519 AA.
AC   P16573;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ecto-ATPase precursor (Cell-CAM 105) (C-CAM 105) (ATP-dependent
DE   taurocolate-carrier protein) (GP110).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 50-68.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX   MEDLINE=89340561; PubMed=2527235;
RA   Lin S.-H., Guidotti G.;
RT   "Cloning and expression of a cDNA coding for a rat liver plasma
RT   membrane ecto-ATPase. The primary structure of the ecto-ATPase is
RT   similar to that of the human biliary glycoprotein I.";
RL   J. Biol. Chem. 264:14408-14414(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY, AND WISTAR; TISSUE=Liver;
RX   MEDLINE=93279310; PubMed=8504806;
RA   Edlund M., Gaardsvoll H., Bock E., Oebrink B.;
RT   "Different isoforms and stock-specific variants of the cell adhesion
RT   molecule C-CAM (cell-CAM 105) in rat liver.";
RL   Eur. J. Biochem. 213:1109-1116(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX   MEDLINE=92344597; PubMed=1637321;
RA   Culic O., Huang Q., Flanagan D., Hixon D., Lin S.-H.;
RT   "Molecular cloning and expression of a new rat liver cell-CAM105
RT   isoform. Differential phosphorylation of isoforms.";
RL   Biochem. J. 285:47-53(1992).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestine;
RX   MEDLINE=94058980; PubMed=8240240;
RA   Cheung P.H., Culic O., Qiu Y., Earley K., Thompson N., Hixson D.C.,
RA   Lin S.-H.;
RT   "The cytoplasmic domain of C-CAM is required for C-CAM-mediated
RT   adhesion function: studies of a C-CAM transcript containing an
RT   unspliced intron.";
RL   Biochem. J. 295:427-435(1993).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=90292222; PubMed=2141577;
RA   Aurivillius M., Hansen O.C., Lazrek M.B.S., Bock E., Oebrink B.;
RT   "The cell adhesion molecule Cell-CAM 105 is an ecto-ATPase and a
RT   member of the immunoglobulin superfamily.";
RL   FEBS Lett. 264:267-269(1990).
RN   [6]
RP   PARTIAL SEQUENCE.
RC   STRAIN=WISTAR; TISSUE=Liver;
RX   MEDLINE=93292517; PubMed=8513803;
RA   Becker A., Lucka L., Kilian C., Kannicht C., Reutter W.;
RT   "Characterisation of the ATP-dependent taurocholate-carrier protein
RT   (gp110) of the hepatocyte canalicular membrane.";
RL   Eur. J. Biochem. 214:539-548(1993).
RN   [7]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=93123234; PubMed=8380406;
RA   Najjae S.M., Accili D., Philippe N., Jernberg J., Margolis R.,
RA   Taylor S.I.;
RT   "pp120/ecto-ATPase, an endogenous substrate of the insulin receptor
RT   tyrosine kinase, is expressed as two variably spliced isoforms.";
RL   J. Biol. Chem. 268:1201-1206(1993).
RN   [8]
RP   CHARACTERIZATION.
RX   MEDLINE=91354197; PubMed=1831973;
RA   Lin S.-H., Culic O., Flanagan D., Hixson D.C.;
```

```
RT   "Immunochemical characterization of two isoforms of rat liver ecto-
RT   ATPase that show an immunological and structural identity with a
RT   glycoprotein cell-adhesion molecule with Mr 105,000.";
RL   Biochem. J. 278:155-161(1991).
CC   -!- FUNCTION: THIS PROTEIN IS A CALCIUM-INDEPENDENT CELL ADHESION
CC       MOLECULE WITH HOMOPHILIC BINDING PROPERTIES. MAY PLAY A ROLE
CC       IN THE FORMATION AND MAINTENANCE OF THE SPECIALIZED MEMBRANE
CC       STRUCTURE OF THE APICAL SURFACE OF THE HEPATOCYTES.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CANALICULAR DOMAIN
CC       OF HEPATOCYTE PLASMA MEMBRANES.
CC   -!- ALTERNATIVE PRODUCTS: TWO VARIANTS ARE PRODUCED BY ALTERNATIVE
CC       SPLICING; THEY DIFFER IN THE LENGTHS OF THEIR CYTOPLASMIC DOMAINS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIA, VESSEL ENDOTHELIA,
CC       LEUKOCYTES AND PLATELETS.
CC   -!- PTM: PHOSPHORYLATED ON SERINE RESIDUES IN BOTH ISOFORMS. ON SERINE
CC       AND TYROSINE RESIDUES IN CYTOPLASMIC REGION OF THE LONG ISOFORM.
CC   -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
CC       SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04963; AAA41104.1; -.
DR   EMBL; X71122; CAA50435.1; -.
DR   EMBL; Z12019; CAA78054.1; -.
DR   EMBL; M92848; AAA16783.1; ALT_TERM.
DR   PIR; A44783; A44783.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003598; Ig_c2.
DR   Pfam; PF00047; ig; 4.
DR   SMART; SM00408; IGc2; 2.
KW   Cell adhesion; Transmembrane; Glycoprotein; Immunoglobulin domain;
KW   Repeat; Signal; Phosphorylation; Alternative splicing.
FT   SIGNAL        1     34       POTENTIAL.
FT   CHAIN        35    519       ECTO-ATPASE.
FT   DOMAIN       35    422       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    423    448       POTENTIAL.
FT   DOMAIN      449    519       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       42    140       IG-LIKE V-TYPE DOMAIN.
FT   DOMAIN      160    222       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN      252    306       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN      337    399       IG-LIKE C2-TYPE DOMAIN 3.
FT   DISULFID    167    215       BY SIMILARITY.
FT   DISULFID    259    299       BY SIMILARITY.
FT   DISULFID    344    392       BY SIMILARITY.
FT   MOD_RES      35     35       PYRROLIDONE CARBOXYLIC ACID
FT                                (BY SIMILARITY).
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    152    152       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    173    173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    256    256       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    302    302       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    315    315       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    374    374       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    455    458       GSDH -> SGSF (IN ISOFORM C-CAM3).
FT   VARSPLIC    459    519       MISSING (IN ISOFORM C-CAM3).
FT   VARIANT      49     49       S -> K.
FT   VARIANT      55     55       T -> A.
FT   VARIANT      70     70       V -> G.
FT   VARIANT      73     76       TGLN -> LNPD.
FT   VARIANT      86     86       S -> D.
FT   VARIANT      88     88       T -> M.
FT   VARIANT      90     90       Q -> K.
FT   VARIANT      92     92       E -> G.
FT   VARIANT      99     99       V -> E.
FT   VARIANT     118    119       GP -> RA.
FT   VARIANT     125    125       I -> F.
FT   VARIANT     127    127       K -> Q.
SQ   SEQUENCE   519 AA;  57263 MW;  04143928D7975776 CRC64;


  Query Match             7.7%;  Score 146.5;  DB 1;  Length 519;
  Best Local Similarity   21.3%;  Pred. No. 0.00085;
  Matches   74;  Conservative   58;  Mismatches  136;  Indels   79;  Gaps   18;

Qy      28 MIEKAKGETAYLPCR----FTLGPEDQGPLDIEW-----LLSPADNQKVDQVIILYSGDK 78
           ::   : :  |  |       |    |   ||: :     ::|| |       | |: |
Db     202 LLNVRRTDKGYYECEARNPATFNRSDPFNLDVIYGPDAPVISPPD-------IYLHQGSN 254

Qy      79 IYDDYYQDLKGRVHF---TSNDLKSGDASINVTNLQLSDIGTYQCKVKK-APGVGNKKIQ 134
           :    : |       :    :  |::    : ::|:  :: ||| | |     |:    ::
Db     255 LNLSCHADSNPPAQYFWLINEKLQTSSQELFISNITTNNSGTYACFVNNTVTGLSRTTVK 314

Qy     135 LTVLLKPSGTRCYVDGSEEIGN-------DFKLKCEPKEGSLPLLYEWQKLSNSQKLPTL 187
           : :|        | :| |           | |  |:  : :   |  | ||| |
Db     315 NITVFEP-----VTQPSIQITNTTVKELGSVTLTCFSKDTGVSV--RW--LFNSQSL--- 362

Qy     188 WLAEMTSPVISVKNAST--------EYSGTYSCTVKN----RVGSDQCLLRLDVVPPSNR 235
              ::|  :   :: ||         | :| | | : |    |:      ::|||:|   :
Db     363 ---QLTDRMTLSQDNSTLRIDPIKREDAGDYQCEISNPVSFRISHP---IKLDVIPDPTQ 416

Qy     236 A------GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTS 289
                 | ||| ||| :  : ||  : :  : ::     :   | |: |  |      |
Db     417 GNSGLSEGAIAGIVIGSVAGVALIAALAYFLYSRKTGGGSD---HRDLTEHKP------S 467

Qy     290 TARSYLGSNHSSLGSMSPSNME--GYSKTQYNQVPSEDFERAPQSPT 334
           |:   || :       ||: ::   ||   :|   |:    | |||
Db     468 TSSHNLGPSDD-----SPNKVDDVSYSVLNFNAQQSKRPTSASSSPT 509


Search completed: May 11, 2002, 13:51:08
Job time: 275 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 11, 2002, 12:48:43 ; Search time 42.66 Seconds
(without alignments)
822.143 Million cell updates/sec

Title: US-09-899-634A-4
Perfect score: 1898
Sequence: 1 MALLLCFVLLCGVADLTRSL.....SRMGAVPVMIPAQSKDGSIV 365

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	197	10.4	299	2	S56749	junctional adhesio
2	164.5	8.7	483	2	T17346	hypothetical prote
3	160.5	8.5	1091	2	A58532	glial cell membran
4	159.5	8.4	4162	2	T42633	connectin/titin -
5	159	8.4	5175	2	T20992	hypothetical prote
6	159	8.4	5198	2	T43290	hemicentin precurs
7	157.5	8.3	309	2	I49503	B-lymphocyte activ
8	152.5	8.0	538	2	I68093	PRR2 delta - human
9	151	8.0	526	1	A32164	biliary glycoprote
10	151	8.0	7962	2	I38346	elastic titin - hu
11	150	7.9	518	2	JC4024	poliovirus recepto
12	148.5	7.8	764	2	A49448	irregular chiasm C
13	148	7.8	526	2	A37821	butyrophilin - bov
14	146.5	7.7	519	2	A44783	ecto-ATPase precur
15	145.5	7.7	509	2	JC5288	SHP substrate-1 pr
16	144.5	7.6	513	2	JC5289	SHP substrate-1 pr
17	139.5	7.3	738	2	A40096	platelet-endotheli
18	137.5	7.2	398	2	I49443	gene 2B4 protein -
19	136.5	7.2	3707	2	S18252	heparan sulfate pr
20	136	7.2	309	2	I49522	gene B7-2 protein
21	134	7.1	523	2	I50478	neurolin - goldfis
22	133	7.0	329	1	A48754	B7-2 antigen - hum
23	133	7.0	526	2	S70587	butyrophilin precu
24	133	7.0	530	2	A53437	poliovirus recepto
25	132	7.0	1070	2	JC4593	protein-tyrosine k
26	131	6.9	1323	2	PN0568	connectin 3B - chi
27	129.5	6.8	478	2	I53960	PRR2 alpha - human
28	129.5	6.8	4391	2	A38096	perlecan precursor
29	127.5	6.7	344	2	I56551	neurotrimin - rat
30	127.5	6.7	503	2	JC5287	SHP substrate-1 pr
31	127	6.7	246	1	A32999	myelin P0 protein
32	126.5	6.7	464	2	C30127	transmembrane carc
33	126	6.6	521	2	S34338	biliary glycoprote
34	126	6.6	824	2	S36439	fibroblast growth
35	125.5	6.6	647	2	B41288	vascular cell adhe
36	125.5	6.6	739	2	A41288	vascular cell adhe
37	125.5	6.6	847	2	JH0371	B-cell adhesion pr
38	124.5	6.6	647	2	A35648	B-cell adhesion pr
39	124.5	6.6	858	1	IJRTNC	neural cell adhesi
40	124	6.5	521	2	JC1508	biliary glycoprote
41	124	6.5	1336	2	I60598	Fit-1 tyrosine kin
42	123.5	6.5	417	2	A44194	poliovirus recepto
43	123.5	6.5	538	2	JC2457	vascular cell adhe
44	123.5	6.5	2222	2	T13924	sdk protein - frui
45	123	6.5	1227	2	T23004	hypothetical prote

ALIGNMENTS

RESULT 1
S56749
junctional adhesion molecule precursor - human
N;Alternate names: F11 platelet antigen; platelet adhesion molecule PAM-1; platelet F
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: A59406; S56749
R;Ozaki, H.; Ishii, K.; Horiuchi, H.; Arai, H.; Kawamoto, T.; Okawa, K.; Iwamatsu, A.
J. Immunol. 163, 553-557, 1999
A;Title: Cutting edge: combined treatment of TNF-alpha and IFN-gamma causes redistrib
A;Reference number: A59406; MUID:99323940; PMID:10395639
A;Accession: A59406
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <OZA>
A;Cross-references: GB:AAD42050; NID:g5326797; PIDN:AAD42050.1
R;Naik, U.P.; Ehrlich, Y.H.; Kornecki, E.
Biochem. J. 310, 155-162, 1995
A;Title: Mechanisms of platelet activation by a stimulatory antibody: cross-linking o
A;Reference number: S56749; MUID:95374438; PMID:7646439
A;Accession: S56749
A;Molecule type: protein
A;Residues: 28-49,'X',51-53;62-73,'E',75-103;123,'F',125-130;'FDKDXTIYLNXY';'LT',206,
A;Note: the order of the peptides other than the amino terminus was not determined
C;Genetics:
A;Gene: JAM
C;Keywords: glycoprotein; phosphoprotein; platelet aggregation; platelet membrane
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-299/Product: junctional adhesion molecule #status predicted <MAT>

Query Match 10.4%; Score 197; DB 2; Length 299;
Best Local Similarity 26.3%; Pred. No. 7.8e-07;
Matches 73; Conservative 43; Mismatches 108; Indels 54; Gaps 14;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWL 58
           : |.:  :||| :|    |::: : |  :   :     | | ::   |       :||
Db      12 LCLFILAILLCSLA--LGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPR------VEWK 63

Qy      59 LSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTY 118
             |  :    :: |: :||   |    : || |    | :|    :||     | |||
Db      64 FDQGDTTR----LVCYN-NKITASY----EDRVTF----LPTGITFKSVTR---EDTGTY 107

Qy     119 QCKVKKAPG--VGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQ 176
           | | :  |   |  |::| ||: ||     :  |  |||   | |  ::|| |  | |
Db     108 TCMVSEEGGNSYGEVKVKLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWF 167

Qy     177 K--------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
           |              |||  :      |:    :|  :     :| |||  :|  |:
Db     168 KDGIVMPTNPKSTRAFSNSSYVLNPTTGELVFDPLSASD-----TGEYSCEARNGYGTPM 222

Qy     223 C--LLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVF 258
```

```
            :|:: |      | |  |   ||: |:|:|::||
Db    223 TSNAVRMEAV--ERNVGVIVAA---VLVTLILLGILVF 255


RESULT   2
T17346
hypothetical protein DKFZp586O1624.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17346
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18727
A;Accession: T17346
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-483 <DUE>
A;Cross-references: EMBL:AL117666
A;Experimental source: adult uterus; clone DKFZp586O1624
C;Genetics:
A;Note: DKFZp586O1624.1

  Query Match             8.7%;  Score 164.5;  DB 2;  Length 483;
  Best Local Similarity   25.3%;  Pred. No. 0.00031;
  Matches   62;  Conservative   41;  Mismatches  117;  Indels   25;  Gaps    7;

Qy     107 VTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKE 166
           :|:::: | | | |  : : |  :    ||||  ||      |    :|    |:|:
Db      49 ITDVKIDDAGVYSCTAQNSAGSISANATLTVLETPSLVVPLEDRVVSVGETVALQCK-AT 107

Qy     167 GSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLR 226
           |: |    | |       |         : :: |:|   | :| |:| : | :|:::   :
Db     108 GNPPPRITWFKGDRPLSLTERHHLTPDNQLLVVQNVVAEDAGRYTCEMSNTLGTERAHSQ 167

Qy     227 LDVVPPS--NRAGTIAGA-VIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPP 283
           | |:| :   : ||  |   | |: ::||  |:  |   : |::  |  | :     |||
Db     168 LSVLPAAGCRKDGTTVGIFTIAVVSSIVLTSLVWVCIIYQTRKKSEEYSVTNTDETVVPP 227

Qy     284 PKSRTSTARSYLGSNHSSLGSMSPSNMEGYSKT----------QYNQVPSEDFERAPQSP 333
                  |||    || |::|    |   :|          : | |   |    |:
Db     228 ------DVPSYL----SSQGTLS-DRQETVVRTEGGPQANGHIESNGVCPRDASHFPEPD 276

Qy     334 TLPLA 338
           |  :|
Db     277 THSVA 281


RESULT   3
A58532
glial cell membrane glycoprotein LIG-1 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
C;Accession: A58532
R;Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.
J. Biol. Chem. 271, 22522-22527, 1996
A;Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in
A;Reference number: A58532; MUID:96394313
A;Accession: A58532
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1091 <SUZ>
A;Cross-references: GB:D78572; NID:g1545806; PIDN:BAA11416.1; PID:g1545807
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter
F;36-61/Domain: proteoglycan amino-terminal homology <PAH>
F;71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
F;310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
F;334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
F;358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
F;385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
F;409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
F;440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>


  Query Match             8.5%;  Score 160.5;  DB 2;  Length 1091;
  Best Local Similarity   24.6%;  Pred. No. 0.0017;
  Matches   68;  Conservative   43;  Mismatches  124;  Indels   41;  Gaps   11;

Qy      34 GETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHF 93
           | || | |  |  |  |    | |       ||         |     |:   : |:|
Db     615 GTTARLECAATGHPNPQ----IAW-------QK--------DGG---TDFPAARERRMHV 652

Qy      94 TSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEE 153
             :|    |   :|:::: |:| | |  : : |  :    ||||  ||      |
Db     653 MPDD----DVFF-ITDVKIDDMGVYSCTAQNSAGSVSANATLTVLETPSLAVPLEDRVVT 707

Qy     154 IGNDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCT 213
           :|     :|:      | :  | |      |         : :: |:|   : :| |:|
Db     708 VGETVAFQCKATGSPTPRI-TWLKGGRPLSLTERHHFTPGNQLLVVQNVMIDDAGRYTCE 766

Qy     214 VKNRVGSDQCLLRLDVVPP--SNRAGTIAGA-VIGVLLALVLIGLIVFCCHKKRREEKYE 270
           : | :|:::   :| ::|     : ||  |   | |: ::||  |:  |   : |::  |
Db     767 MSNPLGTERAHSQLSILPTPGCRKDGTTVGIFTIAVVCSIVLTSLVWVCIIYQTRKKSEE 826

Qy     271 KEVHHDIREDVPPPKSRTSTARSYLGSNHSSLGSMS 306
           | :     |||          |||    || |::|
Db     827 YSVTNTDETIVPP------DVPSYL----SSQGTLS 852


RESULT   4
T42633
connectin/titin - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42633
R;Yajima, H.; Ohtsuka, H.; Kawamura, Y.; Kume, H.; Murayama, T.; Abe, H.; Kimura, S.;
Biochem. Biophys. Res. Commun. 223, 160-164, 1996
A;Title: A 11.5-kb 5'-terminal cDNA sequence of chicken breast muscle connectin/titin
A;Reference number: Z22221; MUID:96254045
A;Accession: T42633
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4162 <YAJ>
A;Cross-references: EMBL:D83390; NID:g1513029; PIDN:BAA11908.1; PID:g1513030
A;Experimental source: breast muscle
C;Keywords: skeletal muscle


  Query Match             8.4%;  Score 159.5;  DB 2;  Length 4162;
  Best Local Similarity   25.6%;  Pred. No. 0.011;
  Matches   61;  Conservative   38;  Mismatches   90;  Indels   49;  Gaps   11;

Qy      12 GVADLTRSLSITTPEQMIEK------AKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQ 65
           |  :    |::  | . ::||       |::  | |     ||    |   |
Db    3543 GFQECFAKLTVLEPAVIVEKPGPVKVTAGDSCTLECTVDGTPE----LTARW-------- 3590

Qy      66 KVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKA 125
                  |:::  |:    | :: | :       : : : |  | | | |  :|| :
Db    3591 -------FKDGNELSTDH----KYKISFFNK-----VSGLKILNAGLEDSGEYTFEVKNS 3634

Qy     126 PGVGNKKIQLTV---LLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEW----QKL 178
           |  :    | |   :: || ||   :   ::|:   |:|:   || |:|  |    |::
Db    3635 VGKSSCTASLQVSDRIMPPSFTRKLKETYGQLGSSAVLECK-VYGSPPILVSWFHDGQEI 3693

Qy     179 SNSQKLPTLWLAEMTSPVISVKNASTEYS--GTYSCTVKNRVGSDQCLLRLDV-VPPS 233
           ::  |    : | :|    |:|    : |  ||||||  |  |||:|   | |  |||
Db    3694 TSGDK----YQATLTDNTCSLKVNGLQESDMGTYSCTATNVAGSDECSAFLSVREPPS 3747
```

```
RESULT   5
T20992
hypothetical protein F15G9.4a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T20992; T24733
R;Sulston, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19355
A;Accession: T20992
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5175 <WIL>
A;Cross-references: EMBL:Z47068; PIDN:CAA87335.1; GSPDB:GN00028; CESP:F15G9.4a
A;Experimental source: clone F15G9
R;Kershaw, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19929
A;Accession: T24733
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5175 <WI2>
A;Cross-references: EMBL:Z47070; PIDN:CAA87344.1; GSPDB:GN00028; CESP:F15G9.4a
A;Experimental source: clone T09B9
C;Genetics:
A;Gene: CESP:F15G9.4a
A;Map position: X
A;Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1
1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4879/1; 4941/1; 5011/1; 5077/

  Query Match             8.4%;  Score 159;  DB 2;  Length 5175;
  Best Local Similarity   25.4%;  Pred. No. 0.015;
  Matches   68;  Conservative   32;  Mismatches   92;  Indels   76;  Gaps   12;

Qy      12 GVADLTRSLSITTPEQMIE------KAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQ 65
           || :|   | : ||  :         || |||  | |  :  |  |    ::|
Db    3845 GVKELEMVLDVFTPPVVSVKSDNPIKALGETITLFCNASGNPYPQ----LKW-------- 3892

Qy      66 KVDQVIILYSGDKIYDD---YYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKV 122
                 |  |:|        |||             | ::: :|: :|:| | |:
Db    3893 -------AKGGSLIFDSPDGARISLKG------------ARLDIPHLKKTDVGDYTCQA 3932

Qy     123 KKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSNSQ 182
             | |     : : ||: |   |  :| |  :          | : || |    | |:  :
Db    3933 LNAAGTSEASVSVDVLVPPEINRDGIDMSPRL---------PAQQSLTL----QCLAQGK 3979

Qy     183 KLPTL-WLAEMT--------------SPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRL 227
           :| : |    |               |  | : | |    | |:|  :|  |||  :  :
Db    3980 PVPQMRWTLNGTALTHSTPGITVASDSTFIQINNVSLSDKGVYTCYAENVAGSDNLMYNV 4039

Qy     228 DVV--PPSNRAGT---IAG--AVIGVLL 248
          |||  |  :  ||   | |  |||  |:
Db    4040 DVVQAPVISNGGTKQVIEGELAVIECLV 4067


RESULT   6
T43290
hemicentin precursor - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C;Accession: T43290; T20993; T24734
R;Vogel, B.E.; Hedgecock, E.M.
submitted to the EMBL Data Library, June 1998
A;Description: Hemicentin is required for hemidesmosome mediated cell adhesion and germ-
A;Reference number: Z22396
A;Accession: T43290
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5198 <VOG>
A;Cross-references: EMBL:AF074901; PIDN:AAC26792.1
R;Sulston, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19355
A;Accession: T20993
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5198 <WIL>
A;Cross-references: EMBL:Z47068; PIDN:CAA87336.1; GSPDB:GN00028; CESP:F15G9.4b
A;Experimental source: clone F15G9
R;Kershaw, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: Z19929
A;Accession: T24734
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5198 <WI2>
A;Cross-references: EMBL:Z47070; PIDN:CAA87345.1; GSPDB:GN00028; CESP:F15G9.4b
A;Experimental source: clone T09B9
C;Genetics:
A;Gene: him-4; F15G9.4b
A;Map position: X
A;Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184
; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 303
1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4902/1; 4964/1; 5034/1; 51


  Query Match             8.4%;  Score 159;  DB 2;  Length 5198;
  Best Local Similarity   25.4%;  Pred. No. 0.015;
  Matches   68;  Conservative   32;  Mismatches   92;  Indels   76;  Gaps   12;

Qy      12 GVADLTRSLSITTPEQMIE------KAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQ 65
           || :|   | : ||  :         || |||  | |  :  |  |    ::|
Db    3845 GVKELEMVLDVFTPPVVSVKSDNPIKALGETITLFCNASGNPYPQ----LKW-------- 3892

Qy      66 KVDQVIILYSGDKIYDD---YYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKV 122
                 |  |:|        |||             | ::: :|: :|:| | |:
Db    3893 -------AKGGSLIFDSPDGARISLKG------------ARLDIPHLKKTDVGDYTCQA 3932

Qy     123 KKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSNSQ 182
             | |     : : ||: |   |  :| |  :          | : || |    | |:  :
Db    3933 LNAAGTSEASVSVDVLVPPEINRDGIDMSPRL---------PAQQSLTL----QCLAQGK 3979

Qy     183 KLPTL-WLAEMT--------------SPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRL 227
           :| : |    |               |  | : | |    | |:|  :|  |||  :  :
Db    3980 PVPQMRWTLNGTALTHSTPGITVASDSTFIQINNVSLSDKGVYTCYAENVAGSDNLMYNV 4039

Qy     228 DVV--PPSNRAGT---IAG--AVIGVLL 248
          |||  |  :  ||   | |  |||  |:
Db    4040 DVVQAPVISNGGTKQVIEGELAVIECLV 4067


RESULT   7
I49503
B-lymphocyte activation antigen 7 precursor - mouse
N;Alternate names: MB7-2
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C;Accession: I49503; S17291; I49521
R;Selvakumar, A.; White, P.C.; Dupont, B.
Immunogenetics 38, 292-295, 1993
A;Title: Genomic organization of the murine B-lymphocyte activation antigen B7.
A;Reference number: I49503; MUID:93307789
A;Accession: I49503
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-309 <RES>
A;Cross-references: GB:L12589; NID:g293299; PIDN:AAA37240.1; PID:g293301
R;Freeman, G.J.; Gray, G.S.; Gimmi, C.D.; Lombard, D.B.; Zhou, L.J.; White, M.; Finge
J. Exp. Med. 174, 625-631, 1991
A;Title: Structure, expression, and T cell costimulatory activity of the murine homol
A;Reference number: S17291; MUID:91341422
A;Accession: S17291
```

```
A;Molecule type: mRNA
A;Residues: 1-274,'R',279-309 <FRE>
A;Cross-references: EMBL:X60958; NID:g50111; PIDN:CAA43291.1; PID:g50112
R;Inobe, M.; Linsley, P.S.; Ledbetter, J.A.; Nagai, Y.; Tamakoshi, M.; Uede, T.
Biochem. Biophys. Res. Commun. 200, 443-449, 1994
A;Title: Identification of an alternatively spliced form of the murine homologue of B7.
A;Reference number: I49521; MUID:94220123
A;Accession: I49521
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-143,238-274,'R',279-309 <RE2>
A;Cross-references: GB:D16220; NID:g505118; PIDN:BAA03748.1; PID:g994769
C;Genetics:
A;Gene: B7
A;Introns: 37/1; 143/1; 237/1; 275/1
C;Superfamily: B-lymphocyte restricted antigen B7
C;Keywords: alternative splicing


  Query Match             8.3%;  Score 157.5;  DB 2;  Length 309;
  Best Local Similarity   22.0%;  Pred. No. 0.00056;
  Matches   67;  Conservative   61;  Mismatches  116;  Indels   61;  Gaps   14;

Qy       3 LLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPA 62
           |:| ||||  :: ::  :     :: : |:  :   ||||:    ||:    | |
Db      21 LILLFVLLIRLSQVSSDV-----DEQLSKSVKDKVLLPCRYNSPHEDESEDRIYW----- 70

Qy      63 DNQKVDQVII-LYSGD-KIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
             || |:|:: : :|  |:: :|    | |  : :         |: :  | ||| ||| |
Db      71 --QKHDKVVLSVIAGKLKVWPEY----KNRTLYDNTTY-----SLIILGLVLSDRGTYSC 119

Qy     121 KVKKAPGVGNKKIQLTVLLKPS-----GTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEW 175
           |:|    |  :::   |:| |      |    :      :  :: |    |     : |
Db     120 VVQKKER-GTYEVKHLALVKLSIKADFSTPNITESGNPSADTKRITCFASGGFPKPRFSW 178

Qy     176 QKLSNSQKLPTLWLAEMTSPVISVKNAST--EYSGTYSCTVKNRVGSDQCLLRL------ 227
           | | ::|| :       |   :   |:  ::: | : |:|       ||::
Db     179 --LENGRELPGINTTISQDPESELYTISSQLDFNTTRNHTIK-------CLIKYGDAHVS 229

Qy     228 ---------DVVPPSNRAGTIAGAVIGVLLALVLIGLIV--FCCHKKR----REEKYEKE 272
                  :  | |    : ||  | :: :|:| :|:  || |  :    |  :  :|
Db     230 EDFTWEKPPEDPPDSKNTLVLFGAGFGAVITVVVIVVIIKCFCKHNLQGSCFRRNEASRE 289

Qy     273 VHHDI 277
             :: :
Db     290 TNNSL 294


RESULT   8
I68093
PRR2 delta - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jan-2000
C;Accession: I68093
R;Eberle, F.; Dubreuil, P.; Mattei, M.G.; Devilard, E.; Lopez, M.
Gene 159, 267-272, 1995
A;Title: The human PRR2 gene, related to the human poliovirus receptor gene (PVR), is th
A;Reference number: I53960; MUID:95347610
A;Accession: I68093
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-538 <RES>
A;Cross-references: GB:S79172; NID:g1042204; PID:g1042205
C;Genetics:
A;Gene: PRR2delta
C;Superfamily: poliovirus receptor; immunoglobulin homology
F;276-331/Domain: immunoglobulin homology <IMM>


  Query Match             8.0%;  Score 152.5;  DB 2;  Length 538;
  Best Local Similarity   23.7%;  Pred. No. 0.0026;
  Matches   86;  Conservative   51;  Mismatches  155;  Indels   71;  Gaps   17;

Qy      27 QMIEKAKGETAYLPCRFTLGP---------EDQGPLDIEWLLSPADNQKVDQVIILYSGD 77
           ::| | | :         |:  |          | : |  | || |      |  ||     ||
Db     157 RVIAKPKNQAEAQKVTFSQDPTTVALCISKEGRPPARISWLSSLDWEAKETQV----SG- 211

Qy      78 KIYDDYYQDLKGRVHFTSND--LKSGDASINVTNLQLSDIGTYQCKVK----KAPGVGNK 131
                 | | |  ||    : || |           |  |  |||:    : | :
Db     212 --------TLAGTVTVTSRFTLVPSGRA----------DGVTVTCKVEHESFEEPAL--I 251

Qy     132 KIQLTVLLKPSGTRCYVDGSEEIG-NDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLA 190
           : |:|   |  :     | :  :|  |  | |: :    |  |:|   | :   ||  :|
Db     252 PVTLSVRYPPEVSISGYDDNWYLGRTDATLSCDVRSNPEPTGYDWSTTSGT--FPTSAVA 309

Qy     191 EMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGT-----IAGAVIG 245
           : :  ||      : :: |: ||| | ||  :    : |    | ||      | | :|
Db     310 QGSQLVIHA--VDSLFNTTFVCTVTNAVGMGRAEQVIFVRETPNTAGAGATGGIIGGIIA 367

Qy     246 VLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVP---PPKSRTSTARSYLGSNHSSL 302
           ::|  :    :  | ::|:|:  :     :  |  |   ||  :       : |   :|
Db     368 AIIATAVAATGILICRQQRKEQTLQGAEEDEDLEGPPSYKPPTPKAKLEAQEMPSQLFTL 427

Qy     303 GSMSPSNMEGYSKTQY----NQVPSEDFERAPQSPTL-----PLAKVAAPNLSRMGAVPV 353
           |:   | :    || |          ::  |  : |||     ||     |  : :|: |:
Db     428 GASEHSPL----KTPYFDAGASCTEQEMPRYHELPTLEERSGPL----HPGATSLGS-PI 478

Qy     354 MIP 356
             :|
Db     479 PVP 481


RESULT   9
A32164
biliary glycoprotein 1 precursor, splice form a - human
N;Alternate names: transmembrane carcinoembryonic antigen 1 (TM1-CEA); transmembrane
N;Contains: biliary glycoprotein 1, splice form b; biliary glycoprotein 1, splice for
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: A32164; A30127; B30127; A48078; S45664; S65939; A30847; G44476
R;Hinoda, Y.; Neumaier, M.; Hefta, S.A.; Drzeniek, Z.; Wagener, C.; Shively, L.; Heft
Proc. Natl. Acad. Sci. U.S.A. 86, 1668, 1989
A;Reference number: A32164
A;Contents: erratum
A;Accession: A32164
A;Molecule type: mRNA
A;Residues: 1-526 <HIN>
A;Cross-references: GB:J03858; NID:g179439; PIDN:AAA51826.1; PID:g179440
R;Hinoda, Y.; Neumaier, M.; Hefta, S.A.; Drzeniek, Z.; Wagener, C.; Shively, L.; Heft
Proc. Natl. Acad. Sci. U.S.A. 85, 6959-6963, 1988
A;Title: Molecular cloning of a cDNA coding biliary glycoprotein I: Primary structure
A;Reference number: A94206; MUID:88320555
A;Contents: annotation
A;Note: the sequence shown in this reference has been completely corrected in referen
R;Barnett, T.R.; Kretschmer, A.; Austen, D.A.; Goebel, S.J.; Hart, J.T.; Elting, J.J.
J. Cell Biol. 108, 267-276, 1989
A;Title: Carcinoembryonic antigens: alternative splicing accounts for the multiple mR
A;Reference number: A92752; MUID:89139550
A;Accession: A30127
A;Molecule type: mRNA
A;Residues: 1-526 <BAR1>
A;Cross-references: EMBL:X16354; NID:g37197; PIDN:CAA34404.1; PID:g37198; EMBL:X14784
A;Experimental source: splice form a
A;Accession: B30127
A;Molecule type: mRNA
A;Residues: 1-319,'D',417-526 <BAR2>
A;Cross-references: EMBL:X14831; NID:g37199; PIDN:CAA32940.1; PID:g37200; EMBL:X14784
A;Experimental source: splice form b
R;Barnett, T.R.; Drake, L.; Pickle II, W.
Mol. Cell. Biol. 13, 1273-1282, 1993
A;Title: Human biliary glycoprotein gene: characterization of a family of novel alter
A;Reference number: A48078; MUID:93140765
A;Accession: A48078
A;Molecule type: mRNA
A;Residues: 124-141,'H',417-526 <BAR3>
A;Cross-references: GB:M76742; NID:g179480; PIDN:AAA57142.1; PID:g179481
```

```
A;Experimental source: splice form x
A;Note: sequence extracted from NCBI backbone (NCBIN:123602, NCBIP:123606)
A;Note: neither the complete nucleic acid sequence nor the complete translation are show
R;Hauck, W.; Nedellec, P.; Turbide, C.; Stanners, C.P.; Barnett, T.R.; Beauchemin, N.
Eur. J. Biochem. 223, 529-541, 1994
A;Title: Transcriptional control of the human biliary glycoprotein gene, a CEA gene fami
A;Reference number: S45664; MUID:94333343
A;Accession: S45664
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-21 <HAU>
A;Cross-references: EMBL:X67277; NID:g29447; PIDN:CAA47694.1; PID:g606777
R;Nedellec, P.; Turbide, C.; Beauchemin, N.
Eur. J. Biochem. 231, 104-114, 1995
A;Title: Characterization and transcriptional activity of the mouse biliary glycoprotein
A;Reference number: S65939; MUID:95354678
A;Accession: S65939
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-21 <NED>
A;Cross-references: EMBL:X67277; NID:g29447; PIDN:CAA47694.1; PID:g606777
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
A;Note: only a part of the coding sequence is given
R;Khan, W.N.; Fraengsmyr, L.; Teglund, S.; Israelsson, A.; Bremer, K.; Hammarstroem, S.
Genomics 14, 384-390, 1992
A;Title: Identification of three new genes and estimation of the size of the carcinoembr
A;Reference number: A44476; MUID:93052339
A;Contents: annotation; alignment of related sequences
C;Genetics:
A;Gene: GDB:BGP
A;Cross-references: GDB:127992; OMIM:109770
A;Map position: 19q13.2-19q13.2
C;Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin
C;Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein
F;1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
F;1-34/Domain: signal sequence #status predicted <SIG>
F;35-526/Product: biliary glycoprotein 1, splice form a #status predicted <MATA>
F;35-428/Domain: extracellular #status predicted <EXT>
F;35-319,'D',417-526/Product: biliary glycoprotein 1, splice form b #status predicted <M
F;35-141,'H',417-526/Product: biliary glycoprotein 1, splice form x #status predicted <M
F;160-217/Domain: immunoglobulin homology <IMM1>
F;252-301/Domain: immunoglobulin homology <IMM2>
F;341-398/Domain: immunoglobulin homology <IMM3>
F;425-454/Domain: transmembrane #status predicted <TMM>
F;453-526/Domain: intracellular #status predicted <INT>
F;104,111,115,152,182,197,208,224,232,254,274,288,292,302,309,345,351,363,378,405,475/Bi
```

Query Match 8.0%; Score 151; DB 1; Length 526;
Best Local Similarity 22.4%; Pred. No. 0.0032;
Matches 76; Conservative 56; Mismatches 133; Indels 74; Gaps 18;

```
Qy      49 DQGPLDIEWLLSPADNQKVDQVI--ILYSGD----KIYDDYYQ---DLKGRVHFTSN--- 96
           | || : | : :|    : | |   : |  |       | ||:   :|    :  ||
Db     209 DTGPYECE-IQNPVSANRSDPVTLNVTYGPDTPTISPSDTYYRPGANLSLSCYAASNPPA 267

Qy      97 --------DLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLL--------K 140
                   :     : : |: ::: |:| |    :    |:    |:::      |
Db     268 QYSWLINGTFQQSTQELFIPNITVNNSGSYTCHANNSVTGCNRTTVKTIIVTELSPVVAK 327

Qy     141 P--SGTRCYVDGSEEIGNDFKLKCEPKEGSLPL--LYEWQKLSNSQKLPTLWLAEMTSPV 196
           |    ::  | | :: |   | |   :  : :   :: | | :|:::         :
Db     328 PQIKASKTTVTGDKDSVN---LTCSTNDTGISIRWFFKNQSLPSSERMK----LSQGNTT 380

Qy     197 ISVKNASTEYSGTYSCTVKNRVG---SDQCLLRL--DVVPPSN--RAGTIAGAVIGV--L 247
           :|:     | :||| | | | :    ||  :| :  : :|  |    | ||| ||||  |
Db     381 LSINPVKREDAGTYWCEVFNPISKNQSDPIMLNVNYNALPQENGLSPGAIAGIVIGVVAL 440

Qy     248 LALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHSSLGSMSP 307
           :||: : |  |    |      ::    |: |  |              |||:   |  |
Db     441 VALIAVALACFLHFGKTGRASDQR----DLTEHKPSV-----------SNHTQDHSNDP 484

Qy     308 SNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLS 346
```

```
           |      : |:|         | || | :  |:|:|:
Db     485 PN-------KMNEVTYSTLNFEAQQPTQPTS--ASPSLT 514
```

```
RESULT 10
I38346
elastic titin - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C;Accession: I38346
R;Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A;Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A;Reference number: A57430; MUID:96026330
A;Accession: I38346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-7962 <RES>
A;Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g1017427
C;Genetics:
A;Gene: GDB:TTN
A;Cross-references: GDB:127867; OMIM:188840
A;Map position: 2q31-2q31
```

Query Match 8.0%; Score 151; DB 2; Length 7962;
Best Local Similarity 23.9%; Pred. No. 0.099;
Matches 61; Conservative 31; Mismatches 77; Indels 86; Gaps 12;

```
Qy      12 GVADLTRSLSITTP------EQMIEKAKGETAYLPCRFTLGPEDQG--PLDIEWLLSPAD 63
           | |  : || :  |          ||  ||   :| |      | ||  |  : |
Db    4050 GSASSSTSLKVKEPPIFRKKPHPIETLKGADVHLEC------ELQGTPPFHVSW------ 4097

Qy      64 NQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGD----------ASINVTNLQLS 113
                              |:|          :|:||           ||:: |:  :
Db    4098 --------------------YKD--------KRELRSGKKYKIMSENFLTSIHILNVDAA 4129

Qy     114 DIGTYQCKVKKAPGVGNKKIQLTVLLK--PSGTRCYVDGSEEIGNDFKLKCEPKEGSLPL 171
           ||| ||||      ||:     :: ||  |   :   | |  :| : :|:    ||: |:
Db    4130 DIGEYQCKATN--DVGSDTCVGSIALKAPPRFVKKLSDISTVVGKEVQLQ-TTIEGAEPI 4186

Qy     172 LYEWQK-------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRV 218
             | |              :| |: : ||  :       :   ||      | |:| :||
Db    4187 SVVWFKDKGEIVRESDNIWISYSENIATLQFSR-----VEPANA-----GKYTCQIKNDA 4236

Qy     219 GSDQCLLRLDVVPPS 233
           |  :|   | |: |:
Db    4237 GMQECFATLSVLEPA 4251
```

```
RESULT 11
JC4024
poliovirus receptor-related protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 05-Nov-1999
C;Accession: JC4024
R;Lopez, M.; Eberle, F.; Mattei, M.G.; Gabert, J.; Birg, F.; Bardin, F.; Maroc, C.; D
Gene 155, 261-265, 1995
A;Title: Complementary DNA characterization and chromosomal localization of a human g
A;Reference number: JC4024; MUID:95237621
A;Accession: JC4024
A;Molecule type: mRNA
A;Residues: 1-518 <LOP>
A;Cross-references: EMBL:X76400; NID:g732795; PIDN:CAA53980.1; PID:g732796
C;Genetics:
A;Gene: GDB:PVRR1
A;Cross-references: GDB:583951
A;Map position: 11q23-11q24
C;Superfamily: poliovirus receptor; immunoglobulin homology
C;Keywords: glycoprotein; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-518/Product: poliovirus receptor-related protein #status predicted <MAT>
F;356-379/Domain: transmembrane #status predicted <TMM>
```

```
F;36,72,82,139,287,308,333/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             7.9%;  Score 150;  DB 2;  Length 518;
  Best Local Similarity   18.6%;  Pred. No. 0.0037;
  Matches   75;  Conservative   55;  Mismatches   98;  Indels  176;  Gaps   19;

Qy      34 GETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQV------------IILYS---GDK 78
           |    | | |                 :|  : |: ||           : :|:   |
Db      44 GTDVVLHCSFA---------------NPLPSVKITQVTWQKSTNGSKQNVAIYNPSMGVS 88

Qy      79 IYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQ--LT 136
           :   |    : || |         | :| :: |:| | | | |:    |  ||:: |  ||
Db      89 VLAPY----RERVEFLRPSFT--DGTIRLSRLELEDEGVYICEFATFP-TGNRESQLNLT 141

Qy     137 VLLKPSGTRCYVDGSEEI-----GNDFKL---KCEPKEGSLPLLYEWQ-KLSNSQKLP-- 185
           |: ||:    :::|:: :     | | |:    |   | : |: :|    ::|
Db     142 VMAKPTN---WIEGTQAVLRAKKGQDDKVLVATCTSANGKPPSVVSWETRLKGEARVPGD 198

Qy     186 ----------------------------------------TLWLAEMTSPVISVKN---- 201
                                                   :| |     | ::::
Db     199 SGTPMAPVTVISRYRLVPSREAHQQSLACIVNYHMDRFKESLTLNVQYEPEVTIEGFDGN 258

Qy     202 ------------------ASTEY-----------------------------SGTYSCT 213
                           :|||                              :|||
Db     259 WYLQRMDVKLTCKADANPPATEYHWTTLNGSLPKGVEAQNRTLFFKGPINYSLAGTYICE 318

Qy     214 VKNRVGSDQCLLRLDVV-------PPSN--RAGTIAGAVIG-----VLLALVLIGLIVFC 259
             | :|:    : :::        || :  ||| :  |:||     :|| |:::| ||
Db     319 ATNPIGTRSGQVEVNITEFPYTPSPPEHGRRAGPVPTAIIGGVAGSILLVLIVVGGIVVA 378

Qy     260 CHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHSSLG 303
             ::|          |  : |         || :   |: :|  |
Db     379 LRRRR----------HTFKGDY-------STKKHVYGNGYSKAG 405


RESULT  12
A49448
irregular chiasm C-roughest precursor - fruit fly (Drosophila melanogaster)
N;Alternate names: irreC-roughest protein
C;Species: Drosophila melanogaster
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 07-May-1999
C;Accession: A49448; S34129
R;Ramos, R.G.P.; Igloi, G.L.; Lichte, B.; Baumann, U.; Maier, D.; Schneider, T.; Brandst
Genes Dev. 7, 2533-2547, 1993
A;Title: The irregular chiasm C-roughest locus of Drosophila, which affects axonal proje
A;Reference number: A49448; MUID:94102535
A;Accession: A49448
A;Status: preliminary
A;Molecule type: DNA; mRNA
A;Residues: 1-764 <RAM>
A;Cross-references: GB:L11040; NID:g304790; PID:g304791; EMBL:Z21641; NID:g312985; PID:g
C;Genetics:
A;Gene: FlyBase:rst
A;Cross-references: FlyBase:FBgn0003285
C;Keywords: transmembrane protein


  Query Match             7.8%;  Score 148.5;  DB 2;  Length 764;
  Best Local Similarity   18.0%;  Pred. No. 0.0078;
  Matches   71;  Conservative   62;  Mismatches  122;  Indels  139;  Gaps   14;

Qy      15 DLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQVIILY 74
           |:: : |     | :|   |    | |     |:     :| |:  |:|
Db     341 DISYAPSFRQRPQSMEADVGSVVSLTCEVDSNPQP----EIVWIQHPSD---------- 385

Qy      75 SGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCK------------ 121
                        ||  || :|     : :|:|      | | ||
Db     386 --------------RVVGTSTNL-----TFSVSN---ETAGRYYCKANVPGYAEISADA 422

Qy     122 ---VKKAPGVGNKKIQLTVLLKPSGTRCYVDG---------------SEEIGNDFKLKC 162
             :| :| :|::: |  ::   :   |:                   | | |:|: :
Db     423 YVYLKGSPAIGSQRTQYGLVGDTARIECFASSVPRARHVSWTFNGQEISSESGHDYSILV 482

Qy     163 EPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
           :   | :                              : ::::   : | |:||| |  |:|
Db     483 DAVPGGV-----------------------KSTLIIRDSQAYHYGKYNCTVVNDYGNDV 518

Qy     223 CLLRLDVVPPSNRAGTIAG--AVIGVLLALVLIGLIVFCCHKKRR--------EEKYEK- 271
             ::|       :   || | :|:  || | :: ::   | |: :         | :  |
Db     519 AEIQLQAKKSVSLLMTIVGGISVVAFLLVLTILVVVYIKCKKRTKLPPADVISEHQITKN 578

Qy     272 -------------EVHHDIREDV---------------PPPKSRTSTARSYLGSNHSSLG 303
                        : |:: |:                 |||:  |: :     || ||
Db     579 GGVSCKLEPGDRTSNYSDLKVDISGGYVPYGDYSTHYSPPPQYLTTCSTK---SNGSS-- 633

Qy     304 SMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPL 337
           ::  :| :   : |  |  |   :   |: |||:
Db     634 TIMQNNHQNQLQLQQQQQQSH-HQHHTQTTTLPM 666


RESULT  13
A37821
butyrophilin - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 05-Nov-1999
C;Accession: A37821
R;Jack, L.J.W.; Mather, I.H.
J. Biol. Chem. 265, 14481-14486, 1990
A;Title: Cloning and analysis of cDNA encoding bovine butyrophilin, an apical glycopr

A;Reference number: A37821; MUID:90354441
A;Accession: A37821
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-526 <JAC>
A;Cross-references: GB:M35551; NID:g1763685; PIDN:AAB39766.1; PID:g162773
C;Keywords: transmembrane protein


  Query Match             7.8%;  Score 148;  DB 2;  Length 526;
  Best Local Similarity   23.4%;  Pred. No. 0.0053;
  Matches   71;  Conservative   52;  Mismatches  118;  Indels   62;  Gaps   14;

Qy       4 LLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWL---LS 60
           || |:||      :    :  |:: |   || | ||||  :     :| ::: |    :|
Db      13 LLIFILLQLPKLDSAPFDVIGPQEPILAVVGEDAELPCRLSPNVSAKG-MELRWFREKVS 71

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||       | :   | :   :   : :|||    : :  |  :: :  :: || | |:|
Db      72 PA-------VFVSREGQEQEGEEMAEYRGRVSLVEDHIAEGSVAVRIQEVKASDDGEYRC 124

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKC-------EPK------EG 167
             ::         : | |      |:  ::    :   : :|:|       ||:       |
Db     125 FFRQDENYEEAIVHLKV--AALGSDPHISMKVQESGEIQLECTSVGWYPEPQVQWRTHRG 182

Qy     168 SLPLLYEWQKLSNSQKLPTLWLAEMTSPVI----SVKNASTEYSGTYSCTVKN-RVGSDQ 222
                |:  :| |:       |  : : ||    |:||          || ::|  :| ::
Db     183 E-----EFPSMSESRNPDEEGLFTVRASVIIRDSSMKNV--------SCCIRNLLLGQEK 229

Qy     223 CLLRLDV---VP----PSNRAGTIAGAVIGVLLALVLIGLIVFCCH------KKRREEKY 269
             ||   :|    |      :| ||| |:| |: || | |          ::|| |
Db     230 -----DVEVSIPASFFPRLTPWMVAVAVILVVLGLLTIGSIFFTWRLYKERSRQRRNEFS 284

Qy     270 EKE 272
             ||
Db     285 SKE 287


RESULT  14
A44783
ecto-ATPase precursor - rat
N;Alternate names: canalicular bile acid transport protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 23-Jul-1999
C;Accession: A44783; A44410
```

```
R;Lin, S.H.; Guidotti, G.
J. Biol. Chem. 264, 14408-14414, 1989
A;Title: Cloning and expression of a cDNA coding for a rat liver plasma membrane ecto-AT
A;Reference number: A44783; MUID:89340561
A;Accession: A44783
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-519 <LIN>
A;Cross-references: GB:J04963; NID:g203989; PIDN:AAA41104.1; PID:g203990
R;Sippel, C.J.; Suchy, F.J.; Ananthanarayanan, M.; Perlmutter, D.H.
J. Biol. Chem. 268, 2083-2091, 1993
A;Title: The rat liver ecto-ATPase is also a canalicular bile acid transport protein.
A;Reference number: A44410; MUID:93131966
A;Accession: A44410
A;Molecule type: protein
A;Residues: 110-120;122-138;148-150 <SIP>
C;Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin
C;Keywords: glycoprotein; liver; phosphoprotein; transmembrane protein
F;1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
F;160-217/Domain: immunoglobulin homology <IMM1>
F;252-301/Domain: immunoglobulin homology <IMM2>
F;337-394/Domain: immunoglobulin homology <IMM3>

  Query Match             7.7%;  Score 146.5;  DB 2;  Length 519;
  Best Local Similarity   21.3%;  Pred. No. 0.0066;
  Matches   74;  Conservative   58;  Mismatches  136;  Indels   79;  Gaps   18;

Qy      28 MIEKAKGETAYLPCR----FTLGPEDQGPLDIEW-----LLSPADNQKVDQVIILYSGDK 78
           ::   : :  |  |      |    |   ||: :     ::|| |       | |: |
Db     202 LLNVRRTDKGYYECEARNPATFNRSDPFNLDVIYGPDAPVISPPD-------IYLHQGSN 254

Qy      79 IYDDYYQDLKGRVHF---TSNDLKSGDASINVTNLQLSDIGTYQCKVKK-APGVGNKKIQ 134
           :    : |      :    :  |::    : ::|:  :: ||| | |     |:    ::
Db     255 LNLSCHADSNPPAQYFWLINEKLQTSSQELFISNITTNNSGTYACFVNNTVTGLSRTTVK 314

Qy     135 LTVLLKPSGTRCYVDGSEEIGN-------DFKLKCEPKEGSLPLLYEWQKLSNSQKLPTL 187
           : :|         | :| |          | |  |:  : :   |  | ||| |
Db     315 NITVFEP-----VTQPSIQITNTTVKELGSVTLTCFSKDTGVSV--RW--LFNSQSL--- 362

Qy     188 WLAEMTSPVISVKNAST--------EYSGTYSCTVKN----RVGSDQCLLRLDVVPPSNR 235
           ::|  :   :: ||        | :| | | : |    |:      ::|||:|   :
Db     363 ---QLTDRMTLSQDNSTLRIDPIKREDAGDYQCEISNPVSFRISHP---IKLDVIPDPTQ 416

Qy     236 A------GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTS 289
              | ||| ||| :  : ||  : :  : ::    :   | |: |  |      |
Db     417 GNSGLSEGAIAGIVIGSVAGVALIAALAYFLYSRKTGGGSD---HRDLTEHKP------S 467

Qy     290 TARSYLGSNHSSLGSMSPSNME--GYSKTQYNQVPSEDFERAPQSPT 334
           |:   || :       ||: ::   ||   :|   |:    |  |||
Db     468 TSSHNLGPSDD-----SPNKVDDVSYSVLNFNAQQSKRPTSASSSPT 509


RESULT  15
JC5288
SHP substrate-1 protein, 509 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 20-Jun-2000
C;Accession: JC5288
R;Yamao, T.; Matozaki, T.; Amano, K.; Matsuda, Y.; Takahashi, N.; Ochi, F.; Fujioka, Y.;
Biochem. Biophys. Res. Commun. 231, 61-67, 1997
A;Title: Mouse and human SHPS-1: Molecular cloning of cDNAs and chromosomal localizatio
A;Reference number: JC5287; MUID:97223399
A;Contents: Brain
A;Accession: JC5288
A;Molecule type: mRNA
A;Residues: 1-509 <YAM>
A;Cross-references: DDBJ:D87967; NID:g1864012; PIDN:BAA13520.1; PID:g1864013
C;Comment: This protein is a glycosylated receptor-like protein and plays a role in cell
 acts as a docking protein and induce translocation of SHP-2 from the cytosol to the pla
C;Genetics:
A;Gene: shps-1
A;Map position: 2

  Query Match             7.7%;  Score 145.5;  DB 2;  Length 509;
  Best Local Similarity   19.4%;  Pred. No. 0.0076;
  Matches  103;  Conservative   69;  Mismatches  151;  Indels  207;  Gaps   23;

Qy       3 LLLCFVL--LCGVADLT-RSLSITTPEQMIEKAKGETAYLPCRFT----LGPEDQGPLDI 55
           |||| :|   |    :| : | :| ||: :  | |::  | |  |    :||       |
Db      14 LLLCLLLSASCFCTGVTGKELKVTQPEKSVSVAAGDSTVLNCTLTSLLPVGP-------I 66

Qy      56 EWLLSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSG--DASINVTNLQLS 113
           :|       ::     :::||       :  :   ||   |:  |    | || ::|:
Db      67 KWYRGVGQSR-----LLIYS-------FTGEHFPRVTNVSDATKRNNMDFSIRISNVTPE 114

Qy     114 DIGTYQC-KVKKAPGVGNKKIQ------LTVLLKPSGTRCYVDGSEEIG---NDFKLKCE 163
           | ||| | | :| |   : :||      : || |||     | |  : |         |:
Db     115 DAGTYYCVKFQKGPSEPDTEIQSGGGTEVYVLAKPSPPE--VSGPADRGIPDQKVNFTCK 172

Qy     164 -----PKEGSLPLLYEWQKLS-----------------------NSQKLPTLWLAEM- 192
                |:  :|    : |:|                        ||  : :  : |:
Db     173 SHGFSPRNITLKWFKDGQELHHLETTVNPSGKNVSYNISSTVRVVLNSMDVHSKVICEVA 232

Qy     193 --------------------TSPVISV------------------------------- 199
                              || : |
Db     233 HITLDRSPLRGIANLSNFIRVSPTVKVTQQSPTSMNQVNLTCRAERFYPEDLQLIWLENG 292

Qy     200 --------KNASTEYSGTYSCTVKNRVGSD--------QCLLRLDVVPPSNRAGTIAGAV 243
                   || :    |||: |    | |          | :: |  |   |  |: |
Db     293 NVSRNDTPKNLTKNTDGTYNYTSLFLVNSSAHREDVVFTCQVKHDQQPAITRNHTVLGLA 352

Qy     244 ------------------------IGVLLALVLIGLI--VFCCHKKRRE---------- 266
                                    :||  ||::: |:  ::    |:::
Db     353 HSSDQGSMQTFPGNNATHNWNVFIGVGVACALLVVLLMAALYLLRIKQKKAKGSTSSTRL 412

Qy     267 ---EKYEKEV-----HHDIRE----DVPPPKSRTSTARSYLGSNHSSLGSMSPSNMEGYS 314
              ||  :|:      :|| :    |:  || :    |:   :||:            |:
Db     413 HEPEKNAREITQIQDTNDINDITYADLNLPKEKKPAPRAPEPNNHTE----------YA 461

Qy     315 KTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSKDGSI 364
           :  :||  :        ||  | :   :|||    |   |: |:  |:
Db     462 SIETGKVPRPE-------DTLTYADLDMVHLSRAQPAPKPEPSFSEYASV 504


Search completed: May 11, 2002, 13:48:36
Job time: 3593 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 11, 2002, 12:26:48 ; Search time 28.1 Seconds
(without alignments)
317.272 Million cell updates/sec

Title: US-09-899-634A-4
Perfect score: 1898
Sequence: 1 MALLLCFVLLCGVADLTRSL.....SRMGAVPVMIPAQSKDGSIV 365

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1754	92.4	365	2	US-08-979-424-3	Sequence 3, Appli
2	1754	92.4	365	4	US-09-272-496-2	Sequence 2, Appli
3	1740	91.7	365	4	US-08-928-383B-2	Sequence 2, Appli
4	1740	91.7	365	4	US-08-928-383B-23	Sequence 23, Appl
5	1739	91.6	365	4	US-08-928-383B-26	Sequence 26, Appl
6	1661	87.5	365	4	US-08-928-383B-24	Sequence 24, Appl
7	523	27.6	106	4	US-08-928-383B-8	Sequence 8, Appli
8	453	23.9	95	4	US-08-928-383B-17	Sequence 17, Appl
9	368.5	19.4	319	1	US-08-597-495B-22	Sequence 22, Appl
10	368.5	19.4	319	4	US-09-068-051A-22	Sequence 22, Appl
11	330.5	17.4	318	4	US-09-068-051A-32	Sequence 32, Appl
12	325.5	17.1	390	2	US-08-979-424-1	Sequence 1, Appli
13	276	14.5	387	4	US-09-175-928-2	Sequence 2, Appli
14	197	10.4	299	4	US-09-188-930-331	Sequence 331, App
15	197	10.4	299	4	US-09-462-270-2	Sequence 2, Appli
16	192	10.1	299	4	US-09-188-930-189	Sequence 189, App
17	164.5	8.7	1101	3	US-08-986-485-2	Sequence 2, Appli
18	163.5	8.6	269	4	US-09-430-503-6	Sequence 6, Appli
19	163.5	8.6	269	4	US-09-430-503-8	Sequence 8, Appli
20	161.5	8.5	269	4	US-09-430-503-4	Sequence 4, Appli
21	161.5	8.5	306	4	US-08-205-697A-17	Sequence 17, Appl
22	161.5	8.5	306	4	US-08-702-525-17	Sequence 17, Appl
23	161.5	8.5	306	5	PCT-US95-02576-17	Sequence 17, Appl
24	160.5	8.5	1091	3	US-08-986-485-5	Sequence 5, Appli
25	159.5	8.4	269	4	US-09-430-503-2	Sequence 2, Appli
26	158.5	8.4	306	2	US-08-147-772-4	Sequence 4, Appli
27	158.5	8.4	306	2	US-08-456-104-8	Sequence 8, Appli
28	158.5	8.4	306	2	US-08-101-624-25	Sequence 25, Appl
29	158.5	8.4	306	3	US-08-153-262-4	Sequence 4, Appli
30	158.5	8.4	306	3	US-08-479-744A-31	Sequence 31, Appl
31	158.5	8.4	306	4	US-08-280-757B-31	Sequence 31, Appl
32	158.5	8.4	306	4	US-09-159-135-4	Sequence 4, Appli
33	158.5	8.4	306	4	US-09-450-798-4	Sequence 4, Appli
34	158	8.3	209	4	US-09-430-503-20	Sequence 20, Appl
35	158	8.3	209	4	US-09-430-503-24	Sequence 24, Appl
36	156	8.2	209	4	US-09-430-503-18	Sequence 18, Appl
37	156	8.2	209	4	US-09-430-503-22	Sequence 22, Appl
38	154.5	8.1	320	4	US-08-205-697A-2	Sequence 2, Appli
39	154.5	8.1	320	4	US-08-702-525-2	Sequence 2, Appli
40	154.5	8.1	320	5	PCT-US95-02576-2	Sequence 2, Appli
41	152.5	8.0	581	2	US-08-724-394A-3	Sequence 3, Appli
42	145.5	7.7	581	2	US-08-724-394A-2	Sequence 2, Appli
43	145	7.6	205	4	US-09-462-270-4	Sequence 4, Appli
44	143	7.5	199	4	US-09-430-503-44	Sequence 44, Appl
45	143	7.5	199	4	US-09-430-503-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT   1
US-08-979-424-3
; Sequence 3, Application US/08979424
; Patent No. 5942606
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  VIRAL RECEPTOR PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,424
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0405 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  365 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1946351
US-08-979-424-3


  Query Match             92.4%;  Score 1754;  DB 2;  Length 365;
  Best Local Similarity   90.7%;  Pred. No. 2.5e-155;
```

```
  Matches  331;  Conservative   16;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          ||||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     61 PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 KVKKAPGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| ||||||| ||||||||:||||||||||:|||||||||||:||||||:|| ||
Db    181 SQKMPTSWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||:|||||
Db    241 GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||| ||||||| |||||||||||:|||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAKVAAPNLSRMGAIPVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365


RESULT   2
US-09-272-496-2
; Sequence 2, Application US/09272496
; Patent No. 6245966
; GENERAL INFORMATION:
;  APPLICANT: DeGregori, James
;  TITLE OF INVENTION: Adenoviral mediated gene transfer into lymphocytes
;  FILE REFERENCE: 90-98
;  CURRENT APPLICATION NUMBER: US/09/272,496
;  CURRENT FILING DATE: 1999-03-19
;  EARLIER APPLICATION NUMBER: US 60/092782
;  EARLIER FILING DATE: 1998-07-14
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 365
;   TYPE: PRT
;   ORGANISM: Homo sapiens
09-272-496-2


  Query Match             92.4%;  Score 1754;  DB 4;  Length 365;
  Best Local Similarity   90.7%;  Pred. No. 2.5e-155;
  Matches  331;  Conservative   16;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          |||||||||||||  |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     61 PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 KVKKAPGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| ||||||| ||||||||:||||||||||:|||||||||||:||||||:|| ||
Db    181 SQKMPTSWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||:|||||
Db    241 GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||| ||||||| |||||||||||:|||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAKVAAPNLSRMGAIPVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365


RESULT   3
US-08-928-383B-2
; Sequence 2, Application US/08928383B
; Patent No. 6210921
;  GENERAL INFORMATION:
;    APPLICANT:  Robert W. Finberg, Jeffrey M. Bergelson,
;    APPLICANT:  and Marshall S. Horwitz
;    TITLE OF INVENTION:  CAR, A No. 6210921el Coxsackievirus and Adenovirus
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,383B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,100
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  365 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-928-383B-2


  Query Match             91.7%;  Score 1740;  DB 4;  Length 365;
  Best Local Similarity   90.4%;  Pred. No. 5.1e-154;
  Matches  330;  Conservative   16;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          |||||||||||||  |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     61 PADNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 KVKKAPGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSD 180
```

```
Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:||  |||||| ||||||||:||||||||||:||||||||||||:||||||:|| ||
Db    181 SQKMPTSSLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||:|||||
Db    241 GAIIGTLLALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||| ||||||| ||||||||||||:|||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERTPQSPTLPPAKVAAPNLSRMGAIPVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365
```

RESULT 4
US-08-928-383B-23
; Sequence 23, Application US/08928383B
; Patent No. 6210921
; GENERAL INFORMATION:
; APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
; APPLICANT: and Marshall S. Horwitz
; TITLE OF INVENTION: CAR, A No. 6210921el Coxsackievirus and Adenovirus
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,383B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,100
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-928-383B-23

Query Match 91.7%; Score 1740; DB 4; Length 365;
Best Local Similarity 90.4%; Pred. No. 5.1e-154;
Matches 330; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```
Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          || |||||||||:|| |  |||||||| ||||||||||||:||| ||||||||||||:|
Db      1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |:||| |||||||||||||:|| ||||||||||||:||||||||||||||||||||||
Db     61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| ||| ||||:|||||||:|||||||||||||||||||||||| :||||||:
Db    121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          || :|| |||||||||||||||||:|||||||||:||||||||:||||||||||||||||||
Db    181 SQTMPTSWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||||| |||||||| |:||||:|||||||||||||||||||||||||||||:|||||
Db    241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKVAAPNLSRMGAVPVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365
```

RESULT 5
US-08-928-383B-26
; Sequence 26, Application US/08928383B
; Patent No. 6210921
; GENERAL INFORMATION:
; APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,
; APPLICANT: and Marshall S. Horwitz
; TITLE OF INVENTION: CAR, A No. 6210921el Coxsackievirus and Adenovirus
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,383B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,100
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-383B-26

Query Match 91.6%; Score 1739; DB 4; Length 365;
Best Local Similarity 90.4%; Pred. No. 6.3e-154;
Matches 330; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```
Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
```

```
          || |||||||||:|| |  |||||||| ||||||||||||:||| ||||||||||||:|
Db      1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |:||| ||||||||||||||:|| ||||||||||:|||||||||||||||||||||||
Db     61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||  |||||:|||||||:|||||||||||||||||||||| :||||||:
Db    121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          || :|| ||||||||||||||:||||||||||:|||||||:|||||||||||||||||||
Db    181 SQTMPTPWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||||| |||||||| |:||||:|||||||||||||||||||||||||||||:|||||
Db    241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKVAAPNLSRMGAVPVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365
```

```
RESULT   6
US-08-928-383B-24
; Sequence 24, Application US/08928383B
; Patent No. 6210921
;  GENERAL INFORMATION:
;    APPLICANT:  Robert W. Finberg, Jeffrey M. Bergelson,
;    APPLICANT:  and Marshall S. Horwitz
;    TITLE OF INVENTION:  CAR, A No. 6210921el Coxsackievirus and Adenovirus
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,383B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,100
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  365 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-928-383B-24
```

```
  Query Match           87.5%;  Score 1661;  DB 4;  Length 365;
  Best Local Similarity   87.1%;  Pred. No. 1.2e-146;
  Matches  318;  Conservative   18;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          || |||||||||:|| |  |||||||| ||||||||||||:||| ||||||||||||:|
Db      1 MARLLCFVLLCGIADFTSGLSITTPEQRIEKAKGETAYLPCKFTLSPEDQGPLDIEWLIS 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |:||| ||||||||||||||:|| ||||||||||:|||||||||||||||||||||||
Db     61 PSDNQIVDQVIILYSGDKIYDNYYPDLKGRVHFTSNDVKSGDASINVTNLQLSDIGTYQC 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||  |||||:|||||||:|||||||||||||||||||||| :||||||:
Db    121 KVKKAPGVANKKFLLTVLVKPSGTRCFVDGSEEIGNDFKLKCEPKEGSLPLQFEWQKLSD 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          || :|| ||||||||||||||:||||||||||:|||||||:|||||||||||||||||||
Db    181 SQTMPTSWLAEMTSPVISVKNASSEYSGTYSCTVQNRVGSDQCMLRLDVVPPSNRAGTIA 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||||| |||||||| |:||||:|||||||||||||||||||||||||||||:|||||
Db    241 GAVIGTLLALVLIGAILFCCHRKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHS 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          ||||||||||||||||||||||||||||||||  ||      :    | ||||||||
Db    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLAPAKFKYAYKTDGITVVVMIPAQSK 360

Qy    361 DGSIV 365
          |||||
Db    361 DGSIV 365
```

```
RESULT   7
US-08-928-383B-8
; Sequence 8, Application US/08928383B
; Patent No. 6210921
;  GENERAL INFORMATION:
;    APPLICANT:  Robert W. Finberg, Jeffrey M. Bergelson,
;    APPLICANT:  and Marshall S. Horwitz
;    TITLE OF INVENTION:  CAR, A No. 6210921el Coxsackievirus and Adenovirus
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,383B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,100
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  106 amino acids
```

```
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-928-383B-8


  Query Match             27.6%;  Score 523;  DB 4;  Length 106;
  Best Local Similarity   92.5%;  Pred. No. 1.8e-41;
  Matches   98;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      34 GETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHF 93
           ||||||||:||| |||||||||||||:|||||||||||||||||||||||| ||||||||
Db       1 GETAYLPCKFTLSPEDQGPLDIEWLISPADNQKVDQVIILYSGDKIYDDYYPDLKGRVHF 60

Qy      94 TSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLL 139
           ||||||||||||||||||||||||||||||||||| |||| | ||:
Db      61 TSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVANKKIHLVVLV 106


RESULT   8
US-08-928-383B-17
; Sequence 17, Application US/08928383B
; Patent No. 6210921
;  GENERAL INFORMATION:
;    APPLICANT:  Robert W. Finberg, Jeffrey M. Bergelson,
;    APPLICANT:  and Marshall S. Horwitz
;    TITLE OF INVENTION:  CAR, A No. 6210921el Coxsackievirus and Adenovirus
;    TITLE OF INVENTION:  Receptor
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,383B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,100
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  DFN-020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-928-383B-17


  Query Match             23.9%;  Score 453;  DB 4;  Length 95;
  Best Local Similarity   87.4%;  Pred. No. 5.1e-35;
  Matches   83;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     141 PSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVK 200
           ||| |||||||||||:|||:|||||||||||| |||||||:|||:|| |||||||| |||||
Db       1 PSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPTSWLAEMTSSVISVK 60

Qy     201 NASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNR 235
           |||:||||||||||:||||||||||||:||||||:
Db      61 NASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNK 95


RESULT   9
US-08-597-495B-22
; Sequence 22, Application US/08597495B
; Patent No. 5712369
;  GENERAL INFORMATION:
;    APPLICANT:  Old, Lloyd J.;  Welt, Sydney; Ritter, Gerd;
;    APPLICANT:  Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
;    APPLICANT:  Catimel, B.; Ji, Hong; Burgess, Anthony W.;
;    APPLICANT:  Heath, Joan K.; White, Sara J.; Johnstone, Cameron
;    TITLE OF INVENTION:  Colon Cell And Colon Cancer Cell
;    TITLE OF INVENTION:  Associated Nucleic Acid Molecules, Protein And Peptides
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/597,495B
;      FILING DATE:  02-Feb-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/511,876
;      FILING DATE:  04-Aug-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5712369man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5316.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  319 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-597-495B-22


  Query Match             19.4%;  Score 368.5;  DB 1;  Length 319;
  Best Local Similarity   30.9%;  Pred. No. 2.3e-26;
  Matches   99;  Conservative   62;  Mismatches  128;  Indels   31;  Gaps   13;

Qy      10 LCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW---LLSPADNQK 66
           || |      ::|: ||: ::  ::|::  ||| :      :  | |:|   ||:
Db      12 LCAVRVTVDAISVETPQDVLRASQGKSVTLPCTYHTSTSSREGL-IQWDKLLLT-----H 65

Qy      67 VDQVIILYSGDK--IYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK 124
           ::|:|    :|  |: : |   | ||   ||: :  |||| :  | ::| |||:| |
Db      66 TERVVIWPFSNKNYIHGELY---KNRVSI-SNNAEQSDASITIDQLTMADNGTYECSVSL 121

Qy     125 APGV-GNKK--IQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLS-N 180
           : || |  ::| ||: ||   | ::|   |||: :| |: ||||    | |:: :
Db     122 MSDLEGNTKSRVRLLVLVPPSKPECGIEGETIIGNNIQLTCQSKEGSPTPQYSWKRYNIL 181

Qy     181 SQKLPTLWLAEMTS-PVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTI 239
           :|: |   ||:  |   :|:|| ||: || | ||  |  |:  | : : |  ||
Db     182 NQEQP---LAQPASGQPVSLKNISTDTSGYYICTSSNEEGTQFCNITVAVRSPSMNVALY 238

Qy     240 AGAVIGVLLALVLIGLIVFCC----HKKRREEKYEKEVHHDIREDVPPPKSRTSTAR--- 292
```

```
          |  :||: ||::||:|::||         |:| :   : :  |: |      |  |
Db    239 VGIAVGVVAALIIIGIIIYCCCCRGKDDNTEDKEDARPNREAYEEPPEQLRELSREREEE 298

Qy    293 -SYLGSNHSSLGSMSPSNME 311
           |       | |  || :::
Db    299 DDYRQEEQRSTGRESPDHLD 318


RESULT  10
US-09-068-051A-22
; Sequence 22, Application US/09068051A
; Patent No. 6291235
;   GENERAL INFORMATION:
;        APPLICANT: Old, Lloyd J.;  Welt, Sydney; Ritter, Gerd;
;                   Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
;                   Catimel, B.; Ji, Hong; Burgess, Anthony W.;
;                   Heath, Joan K.; White, Sara J.; Johnstone, Cameron
;        TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
;                            Associated Nucleic Acid Molecules, Protein And Peptides
;        NUMBER OF SEQUENCES: 33
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fulbright & Jaworski LLP
;             STREET: 666 Fifth Avenue
;             CITY: New York City
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10103
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: PC-DOS
;             SOFTWARE: Wordperfect
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/068,051A
;             FILING DATE: 10-Dec-1998
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/597,495
;             FILING DATE: 02-Feb-1996
;             APPLICATION NUMBER: 08/511,876
;             FILING DATE: 04-Aug-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hanson, No. 6291235man D.
;             REGISTRATION NUMBER: 30,946
;             REFERENCE/DOCKET NUMBER: LUD 5316.2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 318-3168
;             TELEFAX: (212) 752-5958
;   INFORMATION FOR SEQ ID NO: 22
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 319 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 22
US-09-068-051A-22


  Query Match           19.4%;  Score 368.5;  DB 4;  Length 319;
  Best Local Similarity   30.9%;  Pred. No. 2.3e-26;
  Matches   99;  Conservative   62;  Mismatches  128;  Indels   31;  Gaps   13;

Qy     10 LCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW---LLSPADNQK 66
          || |     ::|: ||: ::  ::|::  ||| :      :  | |:|   ||:
Db     12 LCAVRVTVDAISVETPQDVLRASQGKSVTLPCTYHTSTSSREGL-IQWDKLLLT-----H 65

Qy     67 VDQVIILYSGDK--IYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK 124
          ::|:|    :|  |: : |   | ||   ||: :  |||| :  | ::| |||:| |
Db     66 TERVVIWPFSNKNYIHGELY---KNRVSI-SNNAEQSDASITIDQLTMADNGTYECSVSL 121

Qy    125 APGV-GNKK--IQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLS-N 180
            : || |  ::| ||: ||   | ::|   |||: :| |: |||    | |:: :
Db    122 MSDLEGNTKSRVRLLVLVPPSKPECGIEGETIIGNNIQLTCQSKEGSPTPQYSWKRYNIL 181

Qy    181 SQKLPTLWLAEMTS-PVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTI 239
          :|: |   ||:  |   :|:|| ||: || | ||  |  |:  | : : |  ||
Db    182 NQEQP---LAQPASGQPVSLKNISTDTSGYYICTSSNEEGTQFCNITVAVRSPSMNVALY 238

Qy    240 AGAVIGVLLALVLIGLIVFCC----HKKRREEKYEKEVHHDIREDVPPPKSRTSTAR--- 292
           |  :||: ||::||:|::||         |:| :   : :  |: |      |  |
Db    239 VGIAVGVVAALIIIGIIIYCCCCRGKDDNTEDKEDARPNREAYEEPPEQLRELSREREEE 298

Qy    293 -SYLGSNHSSLGSMSPSNME 311
           |       | |  || :::
Db    299 DDYRQEEQRSTGRESPDHLD 318


RESULT  11
US-09-068-051A-32
; Sequence 32, Application US/09068051A
; Patent No. 6291235
;   GENERAL INFORMATION:
;        APPLICANT: Old, Lloyd J.;  Welt, Sydney; Ritter, Gerd;
;                   Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
;                   Catimel, B.; Ji, Hong; Burgess, Anthony W.;
;                   Heath, Joan K.; White, Sara J.; Johnstone, Cameron
;        TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
;                            Associated Nucleic Acid Molecules, Protein And Peptides
;        NUMBER OF SEQUENCES: 33
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Fulbright & Jaworski LLP
;             STREET: 666 Fifth Avenue
;             CITY: New York City
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10103
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
;             COMPUTER: IBM PS/2
;             OPERATING SYSTEM: PC-DOS
;             SOFTWARE: Wordperfect
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/068,051A
;             FILING DATE: 10-Dec-1998
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/597,495
;             FILING DATE: 02-Feb-1996
;             APPLICATION NUMBER: 08/511,876
;             FILING DATE: 04-Aug-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hanson, No. 6291235man D.
;             REGISTRATION NUMBER: 30,946
;             REFERENCE/DOCKET NUMBER: LUD 5316.2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 318-3168
;             TELEFAX: (212) 752-5958
;   INFORMATION FOR SEQ ID NO: 32
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 318 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 32
US-09-068-051A-32


  Query Match           17.4%;  Score 330.5;  DB 4;  Length 318;
  Best Local Similarity   30.0%;  Pred. No. 8.1e-23;
  Matches   87;  Conservative   56;  Mismatches  114;  Indels   33;  Gaps   11;

Qy      9 LLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEW--LLSPADNQK 66
          :|| :     :|:: | : ::  |:| :  ||| :     |:    |:|  ||      :
Db     10 MLCAIWVAADALTVETTQDILRAARGRSVTLPCTYNTYVSDREGF-IQWDKLL----RSQ 64

Qy     67 VDQVII--LYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK 124
          ::|:     :   || : |::   ||   ||| :  :||| :  | : | |||:| |
Db     65 TERVVTWNFVTKKYIYGNRYEN---RVR-VSNDAELSNASITIDQLTMDDNGTYECSVSL 120
```

```
Qy    125 APGV---GNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKL-SN 180
                  :::| ||: ||   | : |   |||: :| |   |||    | |:   :
Db    121 MSDQDVNAKSRVRLLVLVPPSKPDCSIQGEMVIGNNIQLTCHSAEGSPSPQYSWKSYNAQ 180

Qy    181 SQKLPTLWLAEMTSPV----ISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRA 236
          :|: |      :| ||    : :|| ||| :| | ||  | || : | : :   |||
Db    181 NQQRP------LTQPVSGEPLLLKNISTETAGYYICTSSNDVGIESCNITVAPRPPSMNI 234

Qy    237 GTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKS 286
            ||    | :||::||:||:||  : :::| :       |||  | :: .
Db    235 ALYAGIAGSVFVALIIIGVIVYCCCCREKDDKDQD------REDARPNRA 278


RESULT  12
US-08-979-424-1
; Sequence 1, Application US/08979424
; Patent No. 5942606
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  VIRAL RECEPTOR PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/979,424
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0405 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  390 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  LUNGFET03
;      CLONE:  1232054
US-08-979-424-1


  Query Match             17.1%;  Score 325.5;  DB 2;  Length 390;
  Best Local Similarity   27.2%;  Pred. No. 3.2e-22;
  Matches  111;  Conservative   65;  Mismatches  157;  Indels   75;  Gaps   16;

Qy      3 LLLCFVLLCGVADLTRS-LSITTPEQMIEKAKGETAYLPCRFTLGPE--DQGPLDIE--- 56
          |   |: |  :|  :|: | :  |   ::  :|    ||  :|| |     | ::
Db     13 LRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSSSQPWEVPFVM 72

Qy     57 WLLSPADNQKVDQVIILYSG--------DKIYDDYYQDLKGRVHFTSNDLKSGDASINVT 108
          |       :| |||:   :|          :|                 : | : |: :
Db     73 WFFK--QKEKEDQVLSYINGVTTSKPGVSLVY----------------SMPSRNLSLRLE 114

Qy    109 NLQLSDIGTYQCKVK------KAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKC 162
           ||  | | | | |        |: |   | ::| ||: |:   | : |   :| : | |
Db    115 GLQEKDSGPYSCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSC 174

Qy    163 EPKEGSLPLLYEWQKLSNSQKLPTLW-----LAEMTSPVISVKNASTEYSGTYSCTVKNR 217
          :       : |:| :     :||:          ::    :|: | |:  :| | |   |
Db    175 QSPRSKPAVQYQWDR-----QLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNE 229

Qy    218 VGSDQCLLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDI 277
          ||: || : |: |      |  :||||:| |: | |:  :|   |   |  |  :|  :||
Db    230 VGTAQCNVTLE-VSTGPGAAVVAGAVVGTLVGLGLLAGLVLLYH---RRGKALEEPANDI 285

Qy    278 RED------VPPPKSRTSTARSYLGSNHSSLGSMSPSNMEGYSKT-QYNQVPSEDFERAP 330
          :||      :| |||  : :::   |: :|  :: | :  |  :       ||   :  |
Db    286 KEDAIAPRTLPWPKSSDTISKNGTLSSVTSARALRPPH--GPPRPGALTPTPSLSSQALP 343

Qy    331 QSPTLPLAKVAAP------------NLSRMGAVPVMIPAQSKDGSIV 365
           || ||    | |              ||||||||||:||||: ||:|
Db    344 -SPRLPTTDGAHPQPISPIPGGVSSSGLSRMGAVPVMVPAQSQAGSLV 390


RESULT  13
US-09-175-928-2
; Sequence 2, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT: McCoy, John M.
;  APPLICANT: LaVallie, Edward R.
;  APPLICANT: Collins-Racie, Lisa A.
;  APPLICANT: Evans, Cheryl
;  APPLICANT: Merberg, David
;  APPLICANT: Treacy, Maurice
;  APPLICANT: Mi, Sha
;  APPLICANT: Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: 6006B.AJ172A
;  CURRENT APPLICATION NUMBER: US/09/175,928A
;  CURRENT FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 387
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-175-928-2


  Query Match             14.5%;  Score 276;  DB 4;  Length 387;
  Best Local Similarity   24.9%;  Pred. No. 1.3e-17;
  Matches   93;  Conservative   62;  Mismatches  155;  Indels   64;  Gaps   13;

Qy      7 FVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQK 66
          |::|  :|     : :| |:  :    |    | | :|     :  | |:|  |     ::
Db      9 FLILSCLAGQVSVVQVTIPDGFVNVTVGSNVTLICIYTTTVASREQLSIQW--SFFHKKE 66

Qy     67 VDQVIILYS--GDKIYDDYYQD-LKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVK 123
          :: : | :|  |  :    ::| : |     |||   |:||| ::::| :| | | | |
Db     67 MEPISIYFSQGGQAVAIGQFKDRITG-----SND--PGNASITISHMQPADSGIYICDVN 119

Qy    124 KAP---GVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
            |   |      : ::||:|||   | | |  | |:   | |     |:   :| | ||
Db    120 NPPDFLGQNQGILNVSVLVKPSKPLCSVQGRPETGHTISLSCLSALGTPSPVYYWHKLEG 179

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
            :|      |: :: : | :    | | ||  ||:|:  |   :|:       | |
Db    180 RDIVPVKENFNPTTGILVIGNLTNFEQGYYQCTAINRLGNSSC--EIDLTSSHPEVGIIV 237

Qy    241 GAVIGVLL-ALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNH 299
          ||:|| |: | ::| :: |  :| : : |                 |:|
Db    238 GALIGSLVGAAIIISVVCFARNKAKAKAKER-------------NSKTI---------- 273
```

```
Qy    300 SSLGSMSPSNMEGYSKTQYNQ--------VPSEDFERAPQS---------PTLPLAKVAA 342
          : |  |:  |  | |:    :          :||   |  | :          ||   |   |
Db    274 AELEPMTKINPRGESEAMPREDATQLEVTLPSSIHETGPDTIQEPDYEPKPTQEPAPEPA 333

Qy    343 PNLSRMGAVPVMIP 356
          |     |: |: :|
Db    334 P-----GSEPMAVP 342


RESULT  14
US-09-188-930-331
; Sequence 331, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT: Strachan, Lorna
;  APPLICANT: Sleeman, Matthew
;  APPLICANT: Onrust, Rene
;  APPLICANT: Murison, James Greg
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION: and Methods For Their Use
;  FILE REFERENCE: 11000.1011c1
;  CURRENT APPLICATION NUMBER: US/09/188,930A
;  CURRENT FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 331
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: Human
US-09-188-930-331


  Query Match             10.4%;  Score 197;  DB 4;  Length 299;
  Best Local Similarity   26.3%;  Pred. No. 2e-10;
  Matches   73;  Conservative   43;  Mismatches  108;  Indels   54;  Gaps   14;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWL 58
          : | :  :||| :|    |::: : |  :   :     | | ::    |       :||
Db     12 LCLFILAILLCSLA--LGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPR------VEWK 63

Qy     59 LSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTY 118
            |  :    :: |: :||   |    : || |    | :|    :||     | |||
Db     64 FDQGDTTR----LVCYN-NKITASY----EDRVTF----LPTGITFKSVTR---EDTGTY 107

Qy    119 QCKVKKAPG--VGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQ 176
           | | :  |   |  |::| ||: ||     :  |  |||   | |  ::|| |  | |
Db    108 TCMVSEEGGNSYGEVKVKLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWF 167

Qy    177 K-------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
          |             |||  :     |:    :|  :     :| |||  :|  |:
Db    168 KDGIVMPTNPKSTRAFSNSSYVLNPTTGELVFDPLSASD-----TGEYSCEARNGYGTPM 222

Qy    223 C--LLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVF 258
             :|:: |      | |  |   ||: |:|:|::||
Db    223 TSNAVRMEAV--ERNVGVIVAA---VLVTLILLGILVF 255


RESULT  15
US-09-462-270-2
; Sequence 2, Application US/09462270
; Patent No. 6358707
; GENERAL INFORMATION:
;  APPLICANT:  SmithKline Beecham Corporation
;  TITLE OF INVENTION:  Human F11 Antigen: A Cell Surface
;  TITLE OF INVENTION:  Receptor   Involved in Platelet Aggregation
;  FILE REFERENCE:  GH-70150US
;  CURRENT APPLICATION NUMBER: US/09/462,270
;  CURRENT FILING DATE:  2000-01-05
;  PRIOR APPLICATION NUMBER:  60/052,186
;  PRIOR FILING DATE:  1997-07-10
;  NUMBER OF SEQ ID NOS:  4
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-462-270-2


  Query Match             10.4%;  Score 197;  DB 4;  Length 299;
  Best Local Similarity   26.3%;  Pred. No. 2e-10;
  Matches   73;  Conservative   43;  Mismatches  108;  Indels   54;  Gaps   14;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFT--LGPEDQGPLDIEWL 58
          : | :  :||| :|    |::: : |  :   :     | | ::    |       :||
Db     12 LCLFILAILLCSLA--LGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPR------VEWK 63

Qy     59 LSPADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTY 118
            |  :    :: |: :||   |    : || |    | :|    :||     | |||
Db     64 FDQGDTTR----LVCYN-NKITASY----EDRVTF----LPTGITFKSVTR---EDTGTY 107

Qy    119 QCKVKKAPG--VGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQ 176
           | | :  |   |  |::| ||: ||     :  |  |||   | |  ::|| |  | |
Db    108 TCMVSEEGGNSYGEVKVKLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWF 167

Qy    177 K-------------LSNSQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQ 222
          |             |||  :     |:    :|  :     :| |||  :|  |:
Db    168 KDGIVMPTNPKSTRAFSNSSYVLNPTTGELVFDPLSASD-----TGEYSCEARNGYGTPM 222

Qy    223 C--LLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVF 258
             :|:: |      | |  |   ||: |:|:|::||
Db    223 TSNAVRMEAV--ERNVGVIVAA---VLVTLILLGILVF 255


Search completed: May 11, 2002, 13:47:22
Job time: 4834 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: May 11, 2002, 12:18:43 ; Search time 55.28 Seconds
(without alignments)
733.392 Million cell updates/sec

Title: US-09-899-634A-4
Perfect score: 1898
Sequence: 1 MALLLCFVLLCGVADLTRSL.....SRMGAVPVMIPAQSKDGSIV 365

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1754	92.4	365	19	AAW69697	Human coxsackievir
2	1754	92.4	365	19	AAW57212	Human coxsackievir
3	1754	92.4	365	22	AAB47270	Human CAR. Homo s
4	1732	91.3	376	19	AAW57213	Mouse coxsackievir
5	1632	86.0	352	22	AAY72878	Human PRO5723 prot
6	1632	86.0	352	22	AAB65294	Human PRO5723 prot
7	1632	86.0	352	22	AAB50930	Human PRO5723 prot
8	1613	85.0	352	19	AAW69698	Mouse coxsackievir
9	1142	60.2	264	20	AAW82729	Adenovirus pACTSG2
10	1142	60.2	277	20	AAW82730	Adenovirus SCAR.RG
11	1142	60.2	397	20	AAW82731	Adenovirus pACSG2S
12	914	48.2	192	22	AAB85732	Rat coxsakie-adeno
13	509.5	26.8	426	22	ABB10359	Human cDNA SEQ ID
14	472	24.9	373	22	AAB48146	Human A236 variant
15	471	24.8	373	20	AAY41692	Human PRO 363 prot
16	471	24.8	373	21	AAB44248	Human PRO363 (UNQ3
17	471	24.8	373	21	AAB33430	Human PRO363 prote
18	471	24.8	373	22	AAU12365	Human PRO363 polyp
19	471	24.8	373	22	AAB48108	Human A236 polypep
20	471	24.8	373	22	AAB48145	Human A236 variant
21	471	24.8	373	22	AAB65293	Human PRO363 prote
22	468	24.7	373	22	AAB48147	Human A236 variant
23	445.5	23.5	373	22	AAB48149	Mouse A236 variant
24	444.5	23.4	373	22	AAB48126	Mouse A236 polypep
25	444.5	23.4	373	22	AAB48148	Mouse A236 variant
26	441.5	23.3	373	22	AAB48150	Mouse A236 variant
27	440.5	23.2	373	22	AAB85862	Murine adipocytes-
28	411	21.7	294	22	AAU18038	Human immunoglobul
29	368.5	19.4	319	18	AAW14146	Human A33 antigen.
30	368.5	19.4	319	20	AAY23323	Amino acid sequenc
31	368.5	19.4	319	22	AAB65863	Human A33 protein
32	361.5	19.0	394	22	AAB65908	Murine secreted pr
33	360.5	19.0	394	22	AAB65840	Murine INTERCEPT 2
34	359.5	18.9	394	22	AAB65911	Murine secreted pr
35	358	18.9	365	22	AAB65841	Murine mature INTE
36	357.5	18.8	273	20	AAY08075	Human A33 protein
37	357.5	18.8	273	20	AAY23329	An A33 related ant
38	357.5	18.8	394	22	AAB65910	Murine secreted pr
39	356.5	18.8	394	22	AAB65909	Murine secreted pr
40	348.5	18.4	259	22	ABB10530	Human cDNA SEQ ID
41	348.5	18.4	259	22	AAU18085	Human immunoglobul
42	347.5	18.3	268	20	AAY08073	Human A33 protein.
43	347.5	18.3	270	20	AAY23327	An A33 related ant
44	330.5	17.4	318	18	AAW14158	Mouse A33 antigen.
45	325.5	17.1	390	20	AAY27096	Human viral recept

ALIGNMENTS

RESULT 1
AAW69697
ID AAW69697 standard; Protein; 365 AA.
XX
AC AAW69697;
XX
DT 07-DEC-1998 (first entry)
XX
DE Human coxsackievirus and Ad2 and Ad5 receptor HCAR protein.
XX
KW HCAR; coxsackievirus receptor; CVB; adenovirus; Ad2 receptor;
KW Ad5 receptor; human; infection; vaccine; therapy.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	Domain	35..130
FT		/note= "extracellular immunoglobulin domain"
FT	Domain	155..220
FT		/note= "extracellular immunoglobulin domain"
FT	Region	45..52
FT		/note= "CVB binding region (Claim 6)"
FT	Region	47..51
FT		/note= "Ad2/5 and CVB binding region (Claim 6)"
FT	Region	53..57
FT		/note= "Ad2/Ad5 binding region (Claim 6)"
FT	Region	69..73
FT		/note= "Ad2/Ad5 binding region (Claim 6)"
FT	Region	72..77
FT		/note= "Ad2/5 and CVB binding region (Claim 6)"
FT	Region	72..77
FT		/note= "CVB-binding region (Claim 6)"
FT	Region	77..79
FT		/note= "Ad2/Ad5 binding region (Claim 6)"

```
FT   Region          96..100
FT                   /note= "CVB-binding region (Claim 6)"
FT   Region          122..127
FT                   /note= "Ad2/5 and CVB binding region (Claim 6)"
XX
PN   WO9833819-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01724.
XX
PR   30-JAN-1997;   97US-0036986.
XX
PA   (UYNY ) UNIV NEW YORK STATE.
XX
PI   Philipson L,  Tomko RP;
XX
DR   WPI; 1998-437397/37.
DR   N-PSDB; AAV50429.
XX
PT   DNA encoding human receptor for adenovirus C and coxsackievirus B -
PT   for preventing and treating viral infection and rendering cells
PT   susceptible to transformation by adenoviral vectors in gene therapy
XX
PS   Claim 3; Page 67-68; 88pp; English.
XX
CC   This is the amino acid sequence of human HCAR, a protein that
CC   serves as a cellular receptor for adenoviruses of the serotypes 2
CC   and 5 (subgroup C) and for the group B coxsackieviruses (CVB).  The
CC   sequence was deduced from an isolated cDNA clone for HCAR (see
CC   AAV50429).  The invention also provides host cells transformed with
CC   DNA molecules encoding HCAR or mouse MCAR (see AAW69698) and methods
CC   of producing the recombinant proteins or their derivatives.  These
CC   proteins, their extracellular domains, as well as oligopeptides
CC   (see AAW69699-708) which bind virus, are claimed.  Isolated HCAR or
CC   MCAR proteins or their fragments or variants are used to prevent or
CC   treat virus infections and for inhibiting the infectivity of Ad2,
CC   Ad5 or CVB.  Methods are also provided for detecting or measuring
CC   the quantity of HCAR or MCAR in a sample, and for identifying
CC   analytes capable of binding to HCAR or MCAR.
XX
SQ   Sequence   365 AA;


  Query Match             92.4%;  Score 1754;  DB 19;  Length 365;
  Best Local Similarity   90.7%;  Pred. No. 6.9e-150;
  Matches  331;  Conservative   16;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
Db       1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
Db      61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
Db     121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
Db     181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
Db     241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
Db     301 slgsmspsnmegysktqynqvpsedfertpqsptlppakvaapnlsrmgaipvmipaqsk 360

Qy     361 DGSIV 365
```

```
Db     361 dgsiv 365


RESULT   2
AAW57212
ID   AAW57212 standard; Protein; 365 AA.
XX
AC   AAW57212;
XX
DT   03-AUG-1998  (first entry)
XX
DE   Human coxsackievirus and adenovirus receptor.
XX
KW   Human; coxsackievirus; adenovirus; receptor; CAR; cardiac infection;
KW   myocarditis; pericarditis; dilated cardiomyopathy; meningoencephalitis;
KW   pancreatic infection; acute pancreatitis; gastrointestinal tract;
KW   diabetes mellitus.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 188
FT                   /note= "encoded by TCG"
XX
PN   WO9811221-A2.
XX
PD   19-MAR-1998.
XX
PF   12-SEP-1997;   97WO-US16189.
XX
PR   13-SEP-1996;   96US-0026100.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Bergelson JM,  Finberg RW,  Horowitz MS;
XX
DR   WPI; 1998-207384/18.
DR   N-PSDB; AAV28845.
XX
PT   DNA encoding coxsackie virus and adenovirus receptor - useful for
PT   modulation of e.g. cardiac, pancreatic or gastrointestinal infection
XX
PS   Claim 12; Fig 1B; 104pp; English.
XX
CC   The present sequence represents human coxsackievirus and adenovirus
CC   receptor (CAR). The present invention also describes: (1) a method for
CC   modulating CAR expression comprising contacting the cell with an agent
CC   which modulates CAR protein activity or CAR nucleic acid expression,
CC   such that a cell associated activity is altered relative to a cell
CC   associated activity of the cell in the absence of the agent; and (2) a
CC   method for detecting the presence of CAR in a biological sample
CC   comprising contacting a biological sample with an agent capable of
CC   detecting CAR protein or mRNA such that the presence of CAR is detected.
CC   Modulation of CAR is useful for treatment of cardiac infection, e.g.
CC   myocarditis, pericarditis or dilated cardiomyopathy, or infection of the
CC   central nervous system, e.g. a non-specific febrile illness or
CC   meningoencephalitis, or pancreatic infection, e.g. acute pancreatitis,
CC   or infection of the respiratory or gastrointestinal tract or childhood
CC   onset diabetes mellitus. Probes derived from CAR nucleic acids are
CC   useful for hybridisation assays, and antibodies raised against CAR
CC   protein are useful for blocking CAR expression. Cell-free assays which
CC   include combining CAR protein and a candidate/test compound are useful
CC   in screening for drugs which interact with CAR protein.
XX
SQ   Sequence   365 AA;


  Query Match             92.4%;  Score 1754;  DB 19;  Length 365;
  Best Local Similarity   90.7%;  Pred. No. 6.9e-150;
  Matches  331;  Conservative   16;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
```

```
          |||||||||||||| |  |||||||||:|||||||||||||||:||| ||||||||||||:|
Db      1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqggpldiewlis 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| ||||||| |||||||:||||||||||:||||||||||||:||||||:|| ||
Db    181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| ||||:||| ||||||||||||||||||||||||||||||||:|||||
Db    241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||| |||||| ||||||||||||:|||||||||
Db    301 slgsmspsnmegysktqynqvpsedfertpqsptlppakvaapnlsrmgaipvmipaqsk 360

Qy    361 DGSIV 365
          |||||
Db    361 dgsiv 365
```

RESULT 3
AAB47270
ID AAB47270 standard; Protein; 365 AA.
XX
AC AAB47270;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human CAR.
XX
KW Coxsackie virus and/or Adenovirus receptor; CAR; transgenic mouse;
KW transgene; tCAR; transduction; integrin; gene transfer; lymphocyte;
KW lymphocyte-specific transcription regulatory sequence; T cell.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..262
FT /note= "Claimed tCAR"
XX
[redacted] US6245966-B1.
[redacted]
PD 12-JUN-2001.
XX
PF 19-MAR-1999; 99US-0272496.
XX
PR 14-JUL-1998; 98US-0092782.
XX
PA (UYTE-) UNIV TECHNOLOGY CORP.
XX
PI DeGregori J;
XX
DR WPI; 2001-366549/38.
DR N-PSDB; AAC85813.
XX
PT Transgenic mouse comprising a nucleotide sequence encoding a truncated
PT Coxsackie virus and/or Adenovirus receptor (tCAR) polypeptide useful for
PT the analysis of cellular physiology -
XX
PS Claim 3; Column 15-18; 21pp; English.
XX
CC This sequence shows a Coxsackie virus and/or Adenovirus receptor (CAR)
CC polypeptide. The DNA encoding this sequence was used in the method of
CC the invention to produce a transgenic mouse whose genome comprises a
CC transgene comprising a nucleotide sequence encoding a truncated CAR
CC polypeptide (tCAR) operably linked to a lymphocyte-specific
CC transcription regulatory nucleotide sequence. The tCAR is expressed
CC in the lymphocytes of the mouse at levels sufficient for increased
CC susceptibility to adenoviral transduction of the lymphocytes without
CC affecting the expression of endogenous integrins in the lymphocytes.
CC The transgenic mouse is useful for the analysis of pathways, via gene
CC transfer, which control various aspects of cellular physiology either
CC in vitro or in vivo. The tCAR-regulatory sequence construct facilitates
CC either the in vivo delivery of genes into T cells by the inoculation of
CC adenovirus recombinants into lymphoid organs or the ex vivo delivery of
CC genes into transgenic T cells as T cells expressing the truncated form
CC of tCAR are highly efficient target cells for adenoviral transduction.
XX
SQ Sequence 365 AA;

Query Match 92.4%; Score 1754; DB 22; Length 365;
Best Local Similarity 90.7%; Pred. No. 6.9e-150;
Matches 331; Conservative 16; Mismatches 18; Indels 0; Gaps 0;

```
Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          . |||||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| ||||||| |||||||:||||||||||:||||||||||||:||||||:|| ||
Db    181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||||:|||||
Db    241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
          |||||||||||||||||||||||||| |||||| ||||||||||||:|||||||||
Db    301 slgsmspsnmegysktqynqvpsedfertpqsptlppakvaapnlsrmgaipvmipaqsk 360

Qy    361 DGSIV 365
          |||||
Db    361 dgsiv 365
```

RESULT 4
AAW57213
ID AAW57213 standard; Protein; 376 AA.
XX
AC AAW57213;
XX
DT 03-AUG-1998 (first entry)
XX
DE Mouse coxsackievirus and adenovirus receptor.
XX
KW Mouse; coxsackievirus; adenovirus; receptor; CAR; cardiac infection;
KW myocarditis; pericarditis; dilated cardiomyopathy; meningoencephalitis;
KW pancreatic infection; acute pancreatitis; gastrointestinal tract;
KW diabetes mellitus.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 366
FT /note= "encoded by TAG, a stop codon, but the
FT sequence is shown to continue"
XX
PN WO9811221-A2.
XX

```
PD   19-MAR-1998.
XX
PF   12-SEP-1997;   97WO-US16189.
XX
PR   13-SEP-1996;   96US-0026100.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Bergelson JM,  Finberg RW,  Horowitz MS;
XX
DR   WPI; 1998-207384/18.
DR   N-PSDB; AAV28846.
XX
PT   DNA encoding coxsackie virus and adenovirus receptor - useful for
PT   modulation of e.g. cardiac, pancreatic or gastrointestinal infection
XX
PS   Disclosure; Fig 14; 104pp; English.
XX
CC   The present sequence represents mouse coxsackievirus and adenovirus
CC   receptor (CAR). The present invention also describes: (1) a method for
CC   modulating CAR expression comprising contacting the cell with an agent
CC   which modulates CAR protein activity or CAR nucleic acid expression,
CC   such that a cell associated activity is altered relative to a cell
CC   associated activity of the cell in the absence of the agent; and (2) a
CC   method for detecting the presence of CAR in a biological sample
CC   comprising contacting a biological sample with an agent capable of
CC   detecting CAR protein or mRNA such that the presence of CAR is detected.
CC   Modulation of CAR is useful for treatment of cardiac infection, e.g.
CC   myocarditis, pericarditis or dilated cardiomyopathy, or infection of the
CC   central nervous system, e.g. a non-specific febrile illness or
CC   meningoencephalitis, or pancreatic infection, e.g. acute pancreatitis,
CC   or infection of the respiratory or gastrointestinal tract or childhood
CC   onset diabetes mellitus. Probes derived from CAR nucleic acids are
CC   useful for hybridisation assays, and antibodies raised against CAR
CC   protein are useful for blocking CAR expression. Cell-free assays which
CC   include combining CAR protein and a candidate/test compound are useful
CC   in screening for drugs which interact with CAR protein.
XX
SQ   Sequence   376 AA;


  Query Match             91.3%;  Score 1732;  DB 19;  Length 376;
  Best Local Similarity   90.1%;  Pred. No. 7e-148;
  Matches  329;  Conservative   17;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           || |||||||||:|| |  |||||||| |||||||||||||:||| ||||||||||||:|
Db       1 marllcfvllcgiadftsglsittpeqriekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |:||| |||||||||||||:|| ||||||||||:|||||||||||||||||||||||
Db      61 psdnqivdqviilysgdkiydnyypdlkgrvhftsndvksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||  |||||:|||||||:|||||||||||||||||||| :||||||:
Db     121 kvkkapgvankkflltvlvkpsgtrcfvdgseeigndfklkcepkegslplqfewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           || :|| |||||||||||||:|||||||||:||||||||:||||||||||||||||||
Db     181 sqtmptpwlaemtspvisvknasseysgtysctvqnrvgsdqcmlrldvvppsnragtia 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           ||||| |||||||| |:|||:| |||||||||||||||||||||||||||:|||||
Db     241 gavigtllalvligailfcchrkgreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQSK 360
           |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db     301 slgsmspsnmegysktqynqvpsedferapqsptlapakvaapnlsrmgavpvmipaqsk 360

Qy     361 DGSIV 365
           |||||
Db     361 dgsiv 365
```

```
RESULT   5
AAY72878
ID   AAY72878 standard; Protein; 352 AA.
XX
AC   AAY72878;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human PRO5723 protein encoded by DNA82361 cDNA clone.
XX
KW   Human; PRO5723; antiinflammatory; dermatological; immunosuppressive;
KW   antirheumatic; antiarthritic; osteopathic; antianaemic; haemostatic;
KW   antithyroid; antidiabetic; antiviral; antipsoriatic; antiallergic;
KW   antiasthmatic; inhibitor; therapy; systemic lupus erythematosis;
KW   spondyloarthropathy; systemic sclerosis; systemic vasculitis;
KW   sarcoidosis; idiopathic inflammatory myopathy; Sjogren's syndrome;
KW   autoimmune thrombocytopenia; immune-mediated renal disease; hepatitis;
KW   demyelinating polyneuropathy; Guillian-Barre syndrome; Whipple's disease;
KW   hepatobiliary disease; primary biliary cirrhosis; sclerosing cholangitis;
KW   inflammatory bowel disease; gluten-sensitive enteropathy; skin disease;
KW   allergic rhinitis; atopic dermatitis; food hypersensitivity; urticaria;
KW   eosinophilic pneumonia; hypersensitivity pneumonitis; graft rejection;
KW   idiopathic pulmonary fibrosis; graft-versus-host-disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide          1..19
FT                    /label= Signal_peptide
FT   Protein          20..352
FT                    /label= Mature_human_PRO5723_protein
FT   Modified-site    31..38
FT                    /label= Tyrosine_kinase_phosphorylation_site
FT   Modified-site    78..84
FT                    /label= Tyrosine_kinase_phosphorylation_site
FT   Region           96..125
FT                    /label= Myelin_PO_protein
FT   Modified-site    106..109
FT                    /note= "Asn is N-glycosylated"
FT   Domain           155..214
FT                    /label= Immunoglobulin_domain
FT   Modified-site    201..204
FT                    /note= "Asn is N-glycosylated"
FT   Region           229..283
FT                    /label= Myelin_PO_protein
FT   Domain           231..256
FT                    /label= Transmembrane_domain
FT   Modified-site    262..269
FT                    /label= Tyrosine_kinase_phosphorylation_site
FT   Modified-site    298..301
FT                    /note= "Asn is N-glycosylated"
XX
PN   WO200116319-A2.
XX
PD   08-MAR-2001.
XX
PF   23-AUG-2000; 2000WO-US23522.
XX
PR   31-AUG-1999;   99US-0151733.
PR   01-SEP-1999;   99WO-US20111.
PR   16-DEC-1999;   99WO-US30095.
PR   18-FEB-2000; 2000WO-US04342.
PR   01-MAR-2000; 2000WO-US05601.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   05-JUN-2000; 2000US-0209832.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Goddard A,  Godowski PJ,  Gurney AL,  Hillan KJ,  Tumas D;
```

PI Watanabe CK, Wood WI;
XX
DR WPI; 2001-226690/23.
DR N-PSDB; AAD02924.
XX
PT New PRO polypeptides for treating immune related and inflammatory
PT diseases such as rheumatoid arthritis, systemic vasculitis, asthma,
PT autoimmune hemolytic anemia, and diabetes mellitus -
XX
PS Claim 10; Fig 10; 118pp; English.
XX
CC The present sequence is PRO5723 protein encoded by DNA82361 cDNA
CC clone. PRO protein, its agonist or antagonist or its antibody which are
CC capable of enhancing or inhibiting the proliferation of T-lymphocytes
CC or of increasing the infiltration of inflammatory cells into a tissue
CC are useful in the diagnosis and treatment of immune-related diseases
CC in mammals. The PRO protein is useful for treating systemic lupus
CC erythematosis, rheumatoid arthritis, osteoarthritis, juvenile chronic
CC arthritis, spondyloarthropathy, systemic sclerosis, idiopathic
CC inflammatory myopathy, Sjogren's syndrome, systemic vasculitis,
CC sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia,
CC thyroiditis, diabetes mellitus, immune-mediated renal disease,
CC demyelinating disease of the central or peripheral nervous system,
CC idiopathic demyelinating polyneuropathy, Guillian-Barre syndrome,
CC chronic inflammatory demyelinating polyneuropathy, hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, autoimmune or immune-mediated skin diseases such as bullous
CC skin disease, erythema multiforme and contact dermatitis, psoriasis,
CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
CC food hypersensitivity and urticaria, immunologic diseases of the lung
CC such as eosinophilic pneumonias, idiopathic pulmonary fibrosis, hyper-
CC sensitivity pneumonitis, transplantation associated diseases such as
CC graft rejection or graft-versus-host-disease.
XX
SQ Sequence 352 AA;

Query Match 86.0%; Score 1632; DB 22; Length 352;
Best Local Similarity 89.5%; Pred. No. 6.9e-139;
Matches 307; Conservative 15; Mismatches 21; Indels 0; Gaps 0;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||||| |  |||||||||:|||||||||||||||:||| ||||||||||||:|
Db       1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| |||||||||||:|||:||||||||||| |||||||:
Db     121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||| |||||||:||||||||||:|||||||||||:||||||:|| ||
Db     181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy     241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
           |||:|| |||| |||||:||| ||||||||||||||||||||||||||||:|||||
Db     241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy     301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAP 343
           |||||||||||||||||||||||||| ||||||| ||   |
Db     301 slgsmspsnmegysktqynqvpsedfertpqsptlppakfkyp 343
```

RESULT 6
AAB65294
ID AAB65294 standard; Protein; 352 AA.
XX
AC AAB65294;
XX
DT 02-APR-2001 (first entry)
XX
DE Human PRO5723 protein sequence SEQ ID NO:505.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay.
XX
OS Homo sapiens.
XX
PN WO200073454-A1.
XX
PD 07-DEC-2000.
XX
PF 30-MAR-2000; 2000WO-US08439.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX
DR WPI; 2001-032160/04.
DR N-PSDB; AAF44263.
XX
PT PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
PS Claim 12; Fig 318; 935pp; English.
XX
CC The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.

```
XX
SQ   Sequence   352 AA;


  Query Match             86.0%;  Score 1632;  DB 22;  Length 352;
  Best Local Similarity   89.5%;  Pred. No. 6.9e-139;
  Matches  307;  Conservative   15;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          |||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| |||||| |||||||:|||||||||:||||||||||:|||||:|| ||
Db    181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||:|| |||| ||||||||| ||||||||||||||||||||||||||||||:|||||
Db    241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAP 343
          |||||||||||||||||||||||||||| |||||||  ||    |
Db    301 slgsmspsnmegysktqynqvpsedfertpqsptlppakfkyp 343


RESULT   7
AAB50930
ID   AAB50930 standard; Protein; 352 AA.
XX
AC   AAB50930;
XX
DT   21-MAR-2001  (first entry)
XX
DE   Human PRO5723 protein.
XX
KW   Human; PRO; antiinflammatory; dermatological; antiarthritic;
KW   antirheumatic; cardiant; antianaemic; immunosuppressive; antithyroid;
KW   antidiabetic; nootropic; neuroprotective; hepatotropic; virucide;
KW   antiallergic; antiasthmatic; immune related disorder;
KW   hepatobiliary disease; autoimmune disease; allergy.
XX
OS   Homo sapiens.
XX
PN   WO200073452-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-US15264.
XX
PR   02-JUN-1999;   99WO-US12252.
PR   20-JUL-1999;   99US-0144732.
PR   20-JUL-1999;   99US-0144758.
PR   28-JUL-1999;   99US-0146222.
PR   01-SEP-1999;   99WO-US20111.
PR   15-SEP-1999;   99WO-US21090.
PR   15-SEP-1999;   99WO-US21547.
PR   29-OCT-1999;   99US-0162506.
PR   30-NOV-1999;   99WO-US28313.
PR   01-DEC-1999;   99WO-US28634.
PR   09-DEC-1999;   99US-0170262.
PR   20-DEC-1999;   99WO-US30911.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   22-FEB-2000; 2000WO-US04414.
PR   24-FEB-2000; 2000WO-US04914.
PR   15-MAR-2000; 2000WO-US06884.
PR   20-MAR-2000; 2000WO-US07377.
PR   21-MAR-2000; 2000WO-US07532.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Chan B,  Goddard A,  Godowski PJ,  Gurney AL;
PI   Hebert C,  Henzel W,  Kabakoff RC,  Shelton DL,  Tumas D,  Watanabe CK;
PI   Wood WI;
XX
DR   WPI; 2001-025253/03.
DR   N-PSDB; AAC91489.
XX
PT   Thirty three nucleic acids encoding PRO polypeptides which are useful
PT   in the diagnosis and treatment of immune related disorders, e.g.
PT   systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
PT   thyroiditis and diabetes mellitus -
XX
PS   Claim 58; Fig 58; 218pp; English.
XX
CC   The present sequence is one of thirty three novel PRO polypeptides.
CC   The PRO polypeptides, anti-PRO antibodies, agonists and
CC   antagonists are useful for treating and diagnosing immune related
CC   disorders such as systemic lupus erythematosus, rheumatoid arthritis,
CC   osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
CC   systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC   syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC   anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
CC   immune-mediated renal disease, demyelinating diseases of the central
CC   and peripheral nervous systems (such as multiple sclerosis, idiopathic
CC   demyelinating polyneuropathy or Guillain-Barre syndrome, and chronic
CC   inflammatory demyelinating polyneuropathy), hepatobiliary diseases
CC   (such as infectious, autoimmune chronic active hepatitis, primary
CC   biliary cirrhosis, granulomatous hepatitis and sclerosing cholangitis),
CC   inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's
CC   disease, autoimmune or immune-mediated skin diseases (such as bullous
CC   skin diseases, erythema multiforme, contact dermatitis, psoriasis),
CC   allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
CC   food hypersensitivity and urticaria), immunological diseases of the
CC   lung (such as eosinophilic pneumonias, idiopathic pulmonary fibrosis
CC   and hypersensitivity pneumonitis), transplantation associated diseases
CC   including graft rejection and graft-versus-host diseases.
XX
SQ   Sequence   352 AA;


  Query Match             86.0%;  Score 1632;  DB 22;  Length 352;
  Best Local Similarity   89.5%;  Pred. No. 6.9e-139;
  Matches  307;  Conservative   15;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          |||||||||||| |  |||||||||:|||||||||||||:||| ||||||||||||:|
Db      1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db    121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          |||:|| |||||| |||||||:|||||||||:||||||||||:|||||:|| ||
Db    181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkaglia 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
```

```
          ||:|| |||| |||||:||| ||||||||||||||||||||||||||||||||:|||||
Db    241 gaiigtllalaligliifccrkkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAKVAAP 343
          |||||||||||||||||||||||||||| ||||||| ||   |
Db    301 slgsmspsnmegysktqynqvpsedfertpqsptlppakfkyp 343


RESULT   8
AAW69698
ID   AAW69698 standard; Protein; 352 AA.
XX
AC   AAW69698;
XX
DT   07-DEC-1998  (first entry)
XX
DE   Mouse coxsackievirus and Ad2 and Ad5 receptor MCAR protein.
XX
KW   MCAR; coxsackievirus receptor; CVB; adenovirus; Ad2 receptor;
KW   Ad5 receptor; mouse; infection; vaccine; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          35..130
FT                   /note= "extracellular immunoglobulin domain"
FT   Domain          155..220
FT                   /note= "extracellular immunoglobulin domain"
FT   Region          45..52
FT                   /note= "CVB binding region (Claim 6)"
FT   Region          47..51
FT                   /note= "Ad2/5 and CVB binding region (Claim 6)"
FT   Region          53..57
FT                   /note= "Ad2/Ad5 binding region (Claim 6)"
FT   Region          69..73
FT                   /note= "Ad2/Ad5 binding region (Claim 6)"
FT   Region          72..77
FT                   /note= "Ad2/5 and CVB binding region (Claim 6)"
FT   Region          72..77
FT                   /note= "CVB-binding region (Claim 6)"
FT   Region          77..79
FT                   /note= "Ad2/Ad5 binding region (Claim 6)"
FT   Region          96..100
FT                   /note= "CVB-binding region (Claim 6)"
FT   Region          122..127
FT                   /note= "Ad2/5 and CVB binding region (Claim 6)"
XX
PN   WO9833819-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US01724.
XX
PR   30-JAN-1997;   97US-0036986.
XX
PA   (UYNY ) UNIV NEW YORK STATE.
XX
PI   Philipson L,  Tomko RP;
XX
DR   WPI; 1998-437397/37.
DR   N-PSDB; AAV50430.
XX
PT   DNA encoding human receptor for adenovirus C and coxsackievirus B -
PT   for preventing and treating viral infection and rendering cells
PT   susceptible to transformation by adenoviral vectors in gene therapy
XX
PS   Claim 3; Page 70-71; 88pp; English.
XX
CC   This is the amino acid sequence of mouse MCAR, a protein that
CC   serves as a cellular receptor for adenoviruses of the serotypes 2
CC   and 5 (subgroup C) and for the group B coxsackieviruses (CVB).  The
CC   sequence was deduced from an isolated cDNA clone for MCAR (see
CC   AAV50430).  The invention also provides host cells transformed with
CC   DNA molecules encoding MCAR or human HCAR (see AAW69697) and methods
CC   of producing the recombinant proteins or their derivatives.  These
CC   proteins, their extracellular domains, as well as oligopeptides
CC   (see AAW69699-708) which bind virus, are claimed.  Isolated HCAR or
CC   MCAR proteins or their fragments or variants are used to prevent or
CC   treat virus infections and for inhibiting the infectivity of Ad2,
CC   Ad5 or CVB.  Methods are also provided for detecting or measuring
CC   the quantity of HCAR or MCAR in a sample, and for identifying
CC   analytes capable of binding to HCAR or MCAR.
XX
SQ   Sequence   352 AA;


  Query Match             85.0%;  Score 1613;  DB 19;  Length 352;
  Best Local Similarity   89.7%;  Pred. No. 3.6e-137;
  Matches  304;  Conservative   17;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
          || |||||||||:|| |  |||||||| |||||||||||||:||| ||||||||||||:|
Db      1 marllcfvllcgiadftsglsittpeqriekakgetaylpckftlspedqgpldiewlis 60

Qy     61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
          |:||| |||||||||||||||:|| ||||||||||||:|||||||||||||||||||||||
Db     61 psdnqivdqviilysgdkiydnyypdlkgrvhftsndvksgdasinvtnlqlsdigtyqc 120

Qy    121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
          |||||||| ||| ||||:|||||||:|||||||||||||||||||||||| :||||||:
Db    121 kvkkapgvankkflltvlvkpsgtrcfvdgseeigndfklkcepkegslplqfewqklsd 180

Qy    181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
          || :|| |||||||||||||||:||||||||:||||||||:|||||||||||||||||||
Db    181 sqtmptpwlaemtspvisvknasseysgtysctvqnrvgsdqcmlrldvvppsnragtia 240

Qy    241 GAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYLGSNHS 300
          ||||| |||||||| |:||||:||||||||||||||||||||||||||||||:|||||
Db    241 gavigtllalvligailfcchrkrreekyekevhhdiredvpppksrtstarsyigsnhs 300

Qy    301 SLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLAK 339
          |||||||||||||||||||||||||||||||||||||  ||
Db    301 slgsmspsnmegysktqynqvpsedferapqsptlapak 339


RESULT   9
AAW82729
ID   AAW82729 standard; Protein; 264 AA.
XX
AC   AAW82729;
XX
DT   29-MAR-1999  (first entry)
XX
DE   Adenovirus pACTSG2-SCAR protein.
XX
KW   pACTSG2-SCAR; chimeric protein; adenoviral fibre protein; monomer;
KW   trimerisation domain; affinity; substrate; gene therapy vector;
KW   attachment; interaction assay; infection.
XX
OS   Mastadenovirus.
OS   Synthetic.
XX
PN   WO9854346-A1.
XX
PD   03-DEC-1998.
XX
PF   28-MAY-1998;   98WO-US11024.
XX
PR   16-JAN-1998;   98US-0071668.
PR   28-MAY-1997;   97US-0047849.
XX
PA   (GENV-) GENVEC INC.
XX
PI   Brough DE,  Einfeld D,  Kovesdi I,  Lizonova A,  Roelvink PW;
PI   Wickham TJ,  Yonehiro G;
XX
```

DR WPI; 1999-059848/05.
DR N-PSDB; AAV72025.
XX
PT New adenoviral fibre trimer with reduced binding to native substrate
PT - useful for, e.g. preparing gene therapy vector with minimal
PT ectopic infection for in vitro applications
XX
PS Example 8; Page 58-59; 103pp; English.
XX
CC This sequence represents a novel adenovirus chimeric protein,
CC pACTSG2-SCAR. This protein is used in a method for the construction of
CC novel monomers having an N-terminus of an adenoviral fibre protein and
CC a trimerisation domain. Such monomers have lower affinity for native
CC substrate than the native adenoviral fibre trimer. Cell lines containing
CC such monomers are used (i) to propagate adenovirus for use as gene
CC therapy vectors (for in vitro or in vivo applications, (ii) as reagents
CC for studying adenoviral attachment and infection, and (iii) in
CC receptor-ligand interaction assays. The new viruses produce minimal
CC ectopic infection (they can not infect native host cells) so are safer as
CC vectors and can be engineered for selective targeting to other cells.
XX
SQ Sequence 264 AA;

Query Match 60.2%; Score 1142; DB 20; Length 264;
Best Local Similarity 88.4%; Pred. No. 8.2e-95;
Matches 214; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||||  |  |||||||||:|||||||||||||||:||| ||||||||||||:|
Db       1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db     121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||| ||||||||:||||||||||:||||||||||:||||||:||: :
Db     181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkagsgs 240

Qy     241 GA 242
           |:
Db     241 gs 242
```

RESULT 10
AAW82730
ID AAW82730 standard; Protein; 277 AA.
XX
AC AAW82730;
XX
DT 29-MAR-1999 (first entry)
XX
DE Adenovirus SCAR.RGD protein.
XX
KW SCAR.RGD; chimeric protein; adenoviral fibre protein; monomer;
KW trimerisation domain; affinity; substrate; gene therapy vector;
KW attachment; interaction assay; infection.
XX
OS Mastadenovirus.
OS Synthetic.
XX
PN WO9854346-A1.
XX
PD 03-DEC-1998.
XX
PF 28-MAY-1998; 98WO-US11024.
XX
PR 16-JAN-1998; 98US-0071668.
PR 28-MAY-1997; 97US-0047849.
XX
PA (GENV-) GENVEC INC.
XX
PI Brough DE, Einfeld D, Kovesdi I, Lizonova A, Roelvink PW;
PI Wickham TJ, Yonehiro G;
XX
DR WPI; 1999-059848/05.
DR N-PSDB; AAV72026.
XX
PT New adenoviral fibre trimer with reduced binding to native substrate
PT - useful for, e.g. preparing gene therapy vector with minimal
PT ectopic infection for in vitro applications
XX
PS Example 8; Page 59-60; 103pp; English.
XX
CC This sequence represents a novel adenovirus chimeric protein, SCAR.RGD.
CC This protein is used in a method for the construction of novel monomers
CC having an N-terminus of an adenoviral fibre protein and a trimerisation
CC domain. Such monomers have lower affinity for native substrate than the
CC native adenoviral fibre trimer. Cell lines containing such monomers are
CC used (i) to propagate adenovirus for use as gene therapy vectors (for in
CC vitro or in vivo applications, (ii) as reagents for studying adenoviral
CC attachment and infection, and (iii) in receptor-ligand interaction
CC assays. The new viruses produce minimal ectopic infection (they can not
CC infect native host cells) so are safer as vectors and can be engineered
CC for selective targeting to other cells.
XX
SQ Sequence 277 AA;

Query Match 60.2%; Score 1142; DB 20; Length 277;
Best Local Similarity 88.4%; Pred. No. 8.8e-95;
Matches 214; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||||  |  |||||||||:|||||||||||||||:||| ||||||||||||:|
Db       1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db     121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||| ||||||||:||||||||||:||||||||||:||||||:||: :
Db     181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkagsgs 240

Qy     241 GA 242
           |:
Db     241 gs 242
```

RESULT 11
AAW82731
ID AAW82731 standard; Protein; 397 AA.
XX
AC AAW82731;
XX
DT 29-MAR-1999 (first entry)
XX
DE Adenovirus pACSG2SCAR.sig chimeric protein.
XX
KW pACSG2SCAR.sigDEL; chimeric protein; adenoviral fibre protein;
KW trimerisation domain; affinity; substrate; gene therapy vector;
KW infection; attachment; interaction assay; vector; monomer.
XX
OS Mastadenovirus.
OS Synthetic.
XX

PN WO9854346-A1.
XX
PD 03-DEC-1998.
XX
PF 28-MAY-1998; 98WO-US11024.
XX
PR 16-JAN-1998; 98US-0071668.
PR 28-MAY-1997; 97US-0047849.
XX
PA (GENV-) GENVEC INC.
XX
PI Brough DE, Einfeld D, Kovesdi I, Lizonova A, Roelvink PW;
PI Wickham TJ, Yonehiro G;
XX
DR WPI; 1999-059848/05.
DR N-PSDB; AAV72027.
XX
PT New adenoviral fibre trimer with reduced binding to native substrate
PT - useful for, e.g. preparing gene therapy vector with minimal
PT ectopic infection for in vitro applications
XX
PS Example 10; Page 60-62; 103pp; English.
XX
CC This sequence represents a novel adenovirus chimeric protein,
CC pACSG2SCAR.sigDEL. This protein is used in a method for the
CC construction of novel monomers having an N-terminus of an adenoviral
CC fibre protein and a trimerisation domain. Such monomers have lower
CC affinity for native substrate than the native adenoviral fibre trimer.
CC Cell lines containing such monomers are used (i) to propagate adenovirus
CC for use as gene therapy vectors (for in vitro or in vivo applications,
CC (ii) as reagents for studying adenoviral attachment and infection, and
CC (iii) in receptor-ligand interaction assays. The new viruses produce
CC minimal ectopic infection (they can not infect native host cells) so are
CC safer as vectors and can be engineered for selective targeting to other
CC cells.
XX
SQ Sequence 397 AA;

Query Match 60.2%; Score 1142; DB 20; Length 397;
Best Local Similarity 88.4%; Pred. No. 1.5e-94;
Matches 214; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |||||||||||||| |  |||||||||:||||||||||||||:||| ||||||||||||:|
Db       1 malllcfvllcgvvdfarslsittpeemiekakgetaylpckftlspedqgpldiewlis 60

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db      61 padnqkvdqviilysgdkiyddyypdlkgrvhftsndlksgdasinvtnlqlsdigtyqc 120

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSN 180
           |||||||| |||| | ||:|||| ||||||||||:|||:|||||||||| |||||||:
Db     121 kvkkapgvankkihlvvlvkpsgarcyvdgseeigsdfkikcepkegslplqyewqklsd 180

Qy     181 SQKLPTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNRAGTIA 240
           |||:|| ||||||| |||||||:|||||||||:||||||||||||:||||||:||: :
Db     181 sqkmptswlaemtssvisvknasseysgtysctvrnrvgsdqcllrlnvvppsnkagsgs 240

Qy     241 GA 242
           |:
Db     241 gs 242
```

RESULT 12
AAB85732
ID AAB85732 standard; Protein; 192 AA.
XX
AC AAB85732;
XX
DT 29-OCT-2001 (first entry)
XX
DE Rat coxsakie-adenovirus receptor (CAR).
XX
KW Sterenzellen; Coxsakie-adenovirus receptor; CAR; liver.
XX
OS Rattus sp.
XX
PN JP2001149095-A.
XX
PD 05-JUN-2001.
XX
PF 24-NOV-1999; 99JP-0333334.
XX
PR 24-NOV-1999; 99JP-0333334.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
PA (HIRO-) ZH HIROSHIMAKEN SANGYO GIJUTSU SHINKO KI.
XX
DR WPI; 2001-499382/55.
DR N-PSDB; AAH76081.
XX
PT Recognition of activated state of liver Sterenzellen, useful for
PT diagnosis of liver disorders -
XX
PS Disclosure; Page 5-6; 8pp; Japanese.
XX
CC The invention relates to a method of recognition of activated state of
CC liver Sterenzellen with expressed amount of Coxsakie-adenovirus receptor
CC (CAR) in a sample containing liver Sterenzellen. A liver disorder can
CC also be diagnosed by comparing expressed amount of CAR in a sample
CC containing liver Sterenzellen with that of normal cells. The method
CC provides a reliable diagnosis of liver disorder with activated liver
CC Sterenzellen. The present sequence represents a rat CAR.
XX
SQ Sequence 192 AA;

Query Match 48.2%; Score 914; DB 22; Length 192;
Best Local Similarity 89.1%; Pred. No. 2e-74;
Matches 171; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

```
Qy      90 RVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPGVGNKKIQLTVLLKPSGTRCYVD 149
           ||||||||:|||||||||||||||||||||||||||||| |:|  ||||:||||||:||
Db       1 rvhftsndvksgdasinvtnlqlsdigtyqckvkkapgvanrkflltvlvkpsgtrcfvd 60

Qy     150 GSEEIGNDFKLKCEPKEGSLPLLYEWQKLSNSQKLPTLWLAEMTSPVISVKNASTEYSGT 209
           || ||||||:|||||||||||| |||||||:|||:|| |||||||||||||||:|||||
Db      61 gsgeigndfrlkcepkegslplqyewqklsdsqkmptpwlaemtspvisvknasseysgt 120

Qy     210 YSCTVKNRVGSDQCLLRLDVVPPSNRAGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKY 269
           ||||::|||||||:|||| |||||||||||||| |||||||| |:|||||||||||
Db     121 ysctiqnrvgsdqcmlrldavppsnragtiagavigtllalvligailfcchkkrreeky 180

Qy     270 EKEVHHDIREDV 281
           ||||||||||||
Db     181 ekevhhdiredv 192
```

RESULT 13
ABB10359
ID ABB10359 standard; Protein; 426 AA.
XX
AC ABB10359;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human cDNA SEQ ID NO: 667.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation.
XX
OS Homo sapiens.
XX

PN WO200154474-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01349.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226868P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-232397P.
PR 14-SEP-2000; 2000US-232398P.
PR 14-SEP-2000; 2000US-232399P.
PR 14-SEP-2000; 2000US-232400P.
PR 14-SEP-2000; 2000US-232401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 13-OCT-2000; 2000US-239937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;

```
XX
DR   WPI; 2001-476161/51.
DR   N-PSDB; ABA06581.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition -
XX
PS   Claim 11; SEQ ID NO: 667; 859pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a protein of the invention.
XX
SQ   Sequence   426 AA;


  Query Match             26.8%;  Score 509.5;  DB 22;  Length 426;
  Best Local Similarity   31.0%;  Pred. No. 2.3e-37;
  Matches  132;  Conservative   72;  Mismatches  149;  Indels   73;  Gaps   12;

Qy       8 VLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLSPADNQKV 67
           :||  :  :  || ::     |: |:|:|| ||| ||        |:: |:::|  |
Db       6 llllslhgvaaslevsespgsiqvargqtavlpctfttsaalin-lnviwmvtplsnanq 64

Qy      68 DQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKKAPG 127
            : :||| | :::|   :   ||| ||   : : : || : | |||| ||||| |   |
Db      65 peqvilyqggqmfdgapr-fhgrvgftgt-mpatnvsifinntqlsdtgtyqclvnnlpd 122

Qy     128 VGNKKI---QLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKLSNSQKL 184
           :| : |    ||||: ||   | : ||::||:|  | |  :||     | |:|| |: ||
Db     123 iggrnigvtgltvlvppsaphcqiqgsqdigsdvillcsseegiprptylwekldntlkl 182

Qy     185 PTLWLAEMTSPVISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNR-AGTIAGAV 243
           |     :    ::::| |   || | |   | :|:  ||| | |: |  |  | ||||:
Db     183 pptatqdqvqgtvtirnisalssglyqcvasnaigtstclldlqvispqprnigliagai 242

Qy     244 -IGVLLALVLIGLIV--FCCHKKRREEKYEKEVHHDIREDVPPPKSRTSTARSY----LG 296
            | :: :  | ||:  |   : : :|: |:|: ::|||  |||   |:|:::
Db     243 gtgaviiifcialilgaffywrsknkeeeeeeipneireddlppk--cssakafhteiss 300

Qy     297 SNHSSLGSMSPSNMEGYS---KTQYN--------------------QVPS--------- 323
           |::::| | :  |   :|   |   |                      :||
Db     301 sdnntltssnaynsrywsnnpkvhrntdsvshfsdlgqsfsfhsgnanipsiyangthlv 360

Qy     324 ------------------------EDFERAPQSPTLPLAKVAAPNLSRMGAVPVMIPAQS 359
                                  | |: |      ::   | |:||||||:||||
Db     361 pgqhktlvvtanrgsspqvmsrsngsvsrkprpphthsytishatlerigavpvmvpaqs 420

Qy     360 KDGSIV 365
           : ||:|
Db     421 ragslv 426


RESULT  14
AAB48146
ID   AAB48146 standard; Protein; 373 AA.
XX
AC   AAB48146;
XX
DT   02-APR-2001  (first entry)
XX
DE   Human A236 variant 2 polypeptide.
XX
KW   TANGO 204; TANGO 206; TANGO 209; A236; secreted protein; human; mouse;
KW   transmembrane protein; antianemic; cerebroprotective; arteriosclerosis;
KW   antiasthmatic; neuroprotective, cytostatic; cardiant; hepatotropic;
KW   antiinflammatory; antidiabetic; antiinfertility; antipyretic; vasotropic;
KW   antirheumatic; nephrotropic; hemostatic; antilipemic; osteopathic;
KW   ophthalmological; antisickling; antiulcer; vulnerary; variant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 28
FT                   /label= E28D
FT                   /note= "wild-type Glu is replaced by Asp"
XX
PN   WO200069885-A2.
XX
PD   23-NOV-2000.
XX
PF   15-MAY-2000; 2000WO-US13361.
XX
PR   14-MAY-1999;   99US-0312359.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Pan Y,  Leiby KR;
XX
DR   WPI; 2001-024999/03.
DR   N-PSDB; AAC84414.
XX
PT   Novel nucleic acids encoding secreted or transmembrane proteins, useful
PT   for treating, e.g. cancer, hemophilia, anemia, ischemia or diseases of
PT   the lung, liver, kidney or pancreas -
XX
PS   Claim 8; Page -; 209pp; English.
XX
CC   The invention provides human and mouse nucleic acids designated TANGO
CC   204, TANGO 206, TANGO 209 and A236 encoding secreted or transmembrane
CC   proteins. The polypeptides, nucleic acids and their modulators may be
CC   useful for treating or modulating cholesterol uptake, blood coagulation,
CC   to modulate cell proliferation, morphogenesis and fate specification,
CC   tissue repair and renewal, to treat cancer and promote wound healing,
CC   modulate angiogenesis, treat hypercholesterolemia, hemophilia, Marfan
CC   syndrome, protein S deficiency, modulate allergic or inflammatory
CC   response, acid secretion, tropic effects on gastrointestinal mucosa, and
CC   promote ulcer healing, treat bone cancer, achandroplasia, myeloma,
CC   fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma, osteoporosis,
CC   leukemia, anemia, thalassemia, cerebral edema, hydrocephalus, brain
CC   herniations, meningitis, ischemic brain or heart disease, infarction,
CC   intracranial hemorage, pancreatitis, diabetes, angina, hypotensive heart
CC   disease, pulmonary heart disease, rheumatic fever, congenital heart
CC   disease, myocardial disease, atherosclerosis, hypertension, jaundice,
CC   hepatic failure, cirrhosis, glomerulonephritis, Goodpasture's syndrome,
CC   sickle cell disease, renal failure, ischemic bowel disease, Crohn's
CC   disease, hernias, hypoadrenalism, hyperadrenalism, Cushing's syndrome,
CC   neoplasia, pulmonary disorders, asthma, ovarian disorders, McCune
CC   Albright syndrome, infertility, uterine disorders, viral disease. The
CC   present sequence represents a human A236 variant polypeptide.
CC   Note: the present variant sequence has been constructed using the
CC   information provided in the specification.
XX
SQ   Sequence   373 AA;


  Query Match             24.9%;  Score 472;  DB 22;  Length 373;
  Best Local Similarity   32.6%;  Pred. No. 4.6e-34;
  Matches  124;  Conservative   70;  Mismatches  154;  Indels   32;  Gaps   10;

Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |:||| :|  |:   :|  |   |::  |  |||  ||  ::  |||||||:
Db       1 mslllllll---vsyyvgtlgthte---ikrvadekvtlpchhqlglpekdtldiewllt 54

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
             ||:   :|:| ||   :|::  :: |||| | || | :|||: :  |: || | | |
Db      55 --dnegnqkvvityssrhvynnlteeqkgrvafasnfl-agdaslqieplkpsdegrytc 111

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKL-- 178
           |||  :       : | ||::||  :| ::|    |:|  |:||   |: |::| ||::
Db     112 kvknsgryvwshvilkvlvrpskpkcelegeltegsdltlqcesssgtepivyywqrire 171

Qy     179 --SNSQKLPTLWLAEMTSP-VISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNR 235
```

```
              ::||     :   |  : ::| :  ||| | ||  |  | : |::|:  |
Db    172 kegederlppksridynhpgrvllqnltmsysglyqctagneagkescvvrv-tvqyvqs 230

Qy    236 AGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVH-HDIREDVPPPKSRTSTARSY 294
          | :|||| |::   :|| |:|:   ::: :|:||:|   ::|||   ||:|     |
Db    231 igmvagavtgivagallifllvwllirrkdkeryeeeerpneiredaeapkarlvkpsss 290

Qy    295 LGSNHSSLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLA-------------KV 340
            : ||    | :     | ::  :  | |   |||      |                ||
Db    291 ssgsrssrsgssstrstansasrsqrtlstd--aapqpglatqayslvgpevrgsepkkv 348

Qy    341 AAPNLSRMGAVPVMIPAQSK 360
             ||::    | |||:||:
Db    349 hhanltkaettpsmipsqsr 368


RESULT 15
AAY41692
ID   AAY41692 standard; Protein; 373 AA.
XX
AC   AAY41692;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO 363 protein sequence.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein.
XX
OS   Homo sapiens.
XX
PN   WO9946281-A2.
XX
PD   16-SEP-1999.
XX
PF   08-MAR-1999;   99WO-US05028.
XX
PR   10-MAR-1998;   98US-0077450.
PR   11-MAR-1998;   98US-0077632.
PR   11-MAR-1998;   98US-0077641.
PR   11-MAR-1998;   98US-0077649.
PR   12-MAR-1998;   98US-0077791.
PR   13-MAR-1998;   98US-0078004.
PR   17-MAR-1998;   98US-0040220.
PR   20-MAR-1998;   98US-0078886.
PR   20-MAR-1998;   98US-0078910.
PR   20-MAR-1998;   98US-0078936.
PR   20-MAR-1998;   98US-0078939.
PR   25-MAR-1998;   98US-0079294.
PR   26-MAR-1998;   98US-0079656.
PR   27-MAR-1998;   98US-0079663.
PR   27-MAR-1998;   98US-0079664.
PR   27-MAR-1998;   98US-0079689.
PR   27-MAR-1998;   98US-0079728.
PR   27-MAR-1998;   98US-0079786.
PR   30-MAR-1998;   98US-0079920.
PR   30-MAR-1998;   98US-0079923.
PR   31-MAR-1998;   98US-0080105.
PR   31-MAR-1998;   98US-0080107.
PR   31-MAR-1998;   98US-0080165.
PR   31-MAR-1998;   98US-0080194.
PR   01-APR-1998;   98US-0080327.
PR   01-APR-1998;   98US-0080328.
PR   01-APR-1998;   98US-0080333.
PR   01-APR-1998;   98US-0080334.
PR   08-APR-1998;   98US-0081049.
PR   08-APR-1998;   98US-0081070.
PR   08-APR-1998;   98US-0081071.
PR   09-APR-1998;   98US-0081195.
PR   09-APR-1998;   98US-0081203.
PR   09-APR-1998;   98US-0081229.
PR   15-APR-1998;   98US-0081817.
PR   15-APR-1998;   98US-0081838.
PR   15-APR-1998;   98US-0081952.
PR   15-APR-1998;   98US-0081955.
PR   21-APR-1998;   98US-0082568.
PR   21-APR-1998;   98US-0082569.
PR   22-APR-1998;   98US-0082700.
PR   22-APR-1998;   98US-0082704.
PR   22-APR-1998;   98US-0082804.
PR   23-APR-1998;   98US-0082767.
PR   23-APR-1998;   98US-0082796.
PR   27-APR-1998;   98US-0083336.
PR   28-APR-1998;   98US-0083322.
PR   29-APR-1998;   98US-0083392.
PR   29-APR-1998;   98US-0083495.
PR   29-APR-1998;   98US-0083496.
PR   29-APR-1998;   98US-0083499.
PR   29-APR-1998;   98US-0083500.
PR   29-APR-1998;   98US-0083545.
PR   29-APR-1998;   98US-0083554.
PR   29-APR-1998;   98US-0083558.
PR   29-APR-1998;   98US-0083559.
PR   30-APR-1998;   98US-0083742.
PR   05-MAY-1998;   98US-0084366.
PR   06-MAY-1998;   98US-0084414.
PR   06-MAY-1998;   98US-0084441.
PR   07-MAY-1998;   98US-0084598.
PR   07-MAY-1998;   98US-0084600.
PR   07-MAY-1998;   98US-0084627.
PR   07-MAY-1998;   98US-0084637.
PR   07-MAY-1998;   98US-0084639.
PR   07-MAY-1998;   98US-0084640.
PR   07-MAY-1998;   98US-0084643.
PR   13-MAY-1998;   98US-0085323.
PR   13-MAY-1998;   98US-0085338.
PR   13-MAY-1998;   98US-0085339.
PR   15-MAY-1998;   98US-0085573.
PR   15-MAY-1998;   98US-0085579.
PR   15-MAY-1998;   98US-0085580.
PR   15-MAY-1998;   98US-0085582.
PR   15-MAY-1998;   98US-0085689.
PR   15-MAY-1998;   98US-0085697.
PR   15-MAY-1998;   98US-0085700.
PR   15-MAY-1998;   98US-0085704.
PR   18-MAY-1998;   98US-0086023.
PR   22-MAY-1998;   98US-0086392.
PR   22-MAY-1998;   98US-0086414.
PR   22-MAY-1998;   98US-0086430.
PR   22-MAY-1998;   98US-0086486.
PR   28-MAY-1998;   98US-0087098.
PR   28-MAY-1998;   98US-0087106.
PR   28-MAY-1998;   98US-0087208.
PR   30-JUL-1998;   98US-0094651.
PR   11-SEP-1998;   98US-0100038.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wood WI,  Goddard A,  Gurney A,  Yuan J,  Baker KP,  Chen J;
XX
DR   WPI; 1999-551358/46.
DR   N-PSDB; AAZ33941.
XX
PT   New secreted and transmembrane polypeptides and their polynucleotides,
PT   useful for treating blood coagulation disorders, cancers and cellular
PT   adhesion disorders -
XX
PS   Claim 12; Fig 24; 530pp; English.
XX
CC   The present invention describes secreted and transmembrane polypeptides
CC   and their polynucleotides. The nucleotide sequences are useful as
CC   sources of probes, primers, for chromosome mapping, and for generation
CC   of antisense sequences. They can also be used to create transgenic
CC   animals. The proteins can be used to treat a variety of diseases and
CC   disorders, depending on their function. Diseases that may be treated
```

CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. AAZ33891 to
CC AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC polypeptide sequence given in the exemplification of the present
CC invention.
XX
SQ Sequence 373 AA;

Query Match 24.8%; Score 471; DB 20; Length 373;
Best Local Similarity 32.6%; Pred. No. 5.6e-34;
Matches 124; Conservative 70; Mismatches 154; Indels 32; Gaps 10;

```
Qy       1 MALLLCFVLLCGVADLTRSLSITTPEQMIEKAKGETAYLPCRFTLGPEDQGPLDIEWLLS 60
           |:|||  :|   |:    :|   |    |::   |   |||   ||  ::  |||||||:
Db       1 mslllllll---vsyyvgtlgthte---ikrvaeekvtlpchhqlglpekdtldiewllt 54

Qy      61 PADNQKVDQVIILYSGDKIYDDYYQDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQC 120
             ||:   :|:| ||   :|::  :: |||| | || | :||||: :  |: || | | |
Db      55 --dnegnqkvvityssrhvynnlteeqkgrvafasnfl-agdaslqieplkpsdegrytc 111

Qy     121 KVKKAPGVGNKKIQLTVLLKPSGTRCYVDGSEEIGNDFKLKCEPKEGSLPLLYEWQKL-- 178
           ||| :       : | ||::||  :| ::|    |:|  |:||   |: |::| ||::
Db     112 kvknsgryvwshvilkvlvrpskpkcelegeltegsdltlqcesssgtepivyywqrire 171

Qy     179 --SNSQKLPTLWLAEMTSP-VISVKNASTEYSGTYSCTVKNRVGSDQCLLRLDVVPPSNR 235
               ::||     :   |  : ::| :  ||| | ||  |  | : |::|:  |
Db     172 kegederlppksridynhpgrvllqnltmsysglyqctagneagkescvvrv-tvqyvqs 230

Qy     236 AGTIAGAVIGVLLALVLIGLIVFCCHKKRREEKYEKEVH-HDIREDVPPPKSRTSTARSY 294
            | :|||| |::   :|| |:|:   ::: :|:||:|   ::|||   ||:|     |
Db     231 igmvagavtgivagallifllvwllirrkdkeryeeeerpneiredaeapkarlvkpsss 290

Qy     295 LGSNHSSLGSMSPSNMEGYSKTQYNQVPSEDFERAPQSPTLPLA----------KV 340
              : ||    | :     | ::  :  | |   |||      |            ||
Db     291 ssgsrssrsgssstrstansasrsqrtlstd--aapqpglatqayslvgpevrgsepkkv 348

Qy     341 AAPNLSRMGAVPVMIPAQSK 360
             ||::    | |||:||:
Db     349 hhanltkaettpsmipsqsr 368
```

Search completed: May 11, 2002, 13:46:26
Job time: 5263 sec